Security of Cyber-Physical Systems

From Theory to Testbeds & Validation

Joaquin Garcia-Alfaro

CNRS SAMOVAR Lab & Télécom SudParis
Université Paris-Saclay

CyberICPS, ESORICS 2016, September 27, 2016

Context

http://www.panoptesec.eu/

Dynamic Risk Approaches for Automated Cyber Defense

FP7-ICT-2013-10

Consortium

RHEA

TELECOM SudParis

INSTITUT Mines-Télécom

s@movar

Supélec

acea

UNIVERSITÄT ZU LÜBECK

SAPIENZA UNIVERSITÀ DI ROMA

epistematica

NOKIA

Context

http://www.panoptesec.eu/

Dynamic Risk Approaches for Automated Cyber Defense

FP7-ICT-2013-10

Consortium

RHEA

TELECOM SudParis

Supélec

acea

UNIVERSITÄT ZU LÜBECK

SAPIENZA UNIVERSITÀ DI ROMA

epistematica

NOKIA

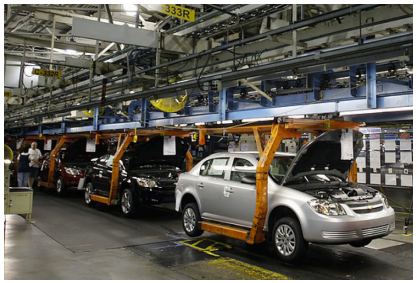
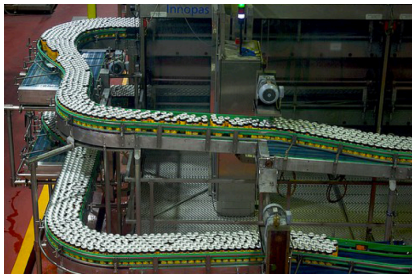
CHAIR OF CYBER CNI Critical National Infrastructures

What is SCADA?

- **Supervisory Control And Data Acquisition**
  - Real-time technologies to centrally monitor/control remote/local equipment
- Distributed Control Systems – E.g., large-scale transmission systems, such as electrical, oil and gas transportation networks

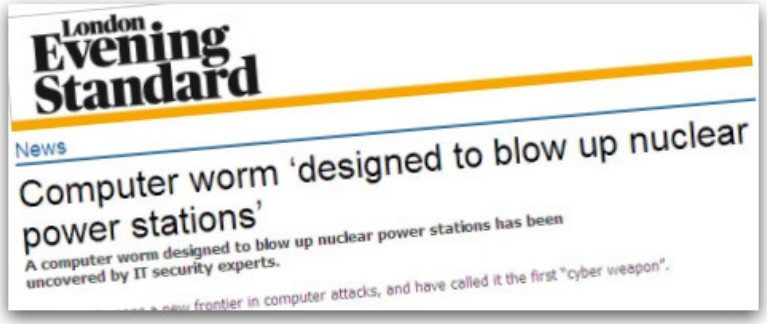


- Industrial Control Systems – Less degree of distribution, but synonym in this talk

Motivation

- Malware moving from IT Systems to Operational Systems
- Wrong configurations, lack of encryption, legacy (vulnerable) systems, third party access, ...


London Evening Standard

News

Computer worm 'designed to blow up nuclear power stations'

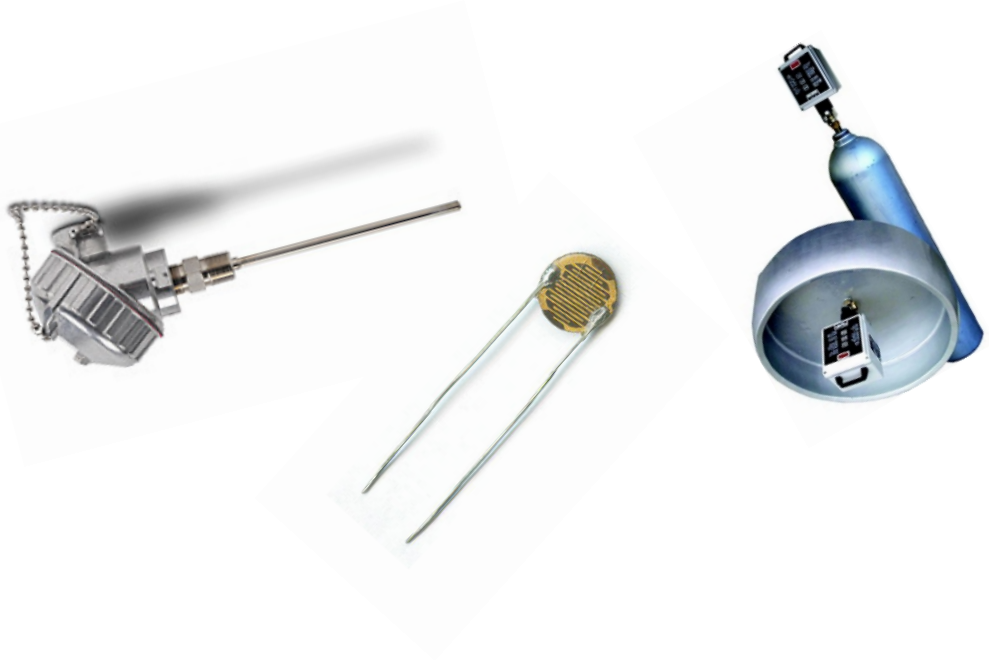
A computer worm designed to blow up nuclear power stations has been uncovered by IT security experts.

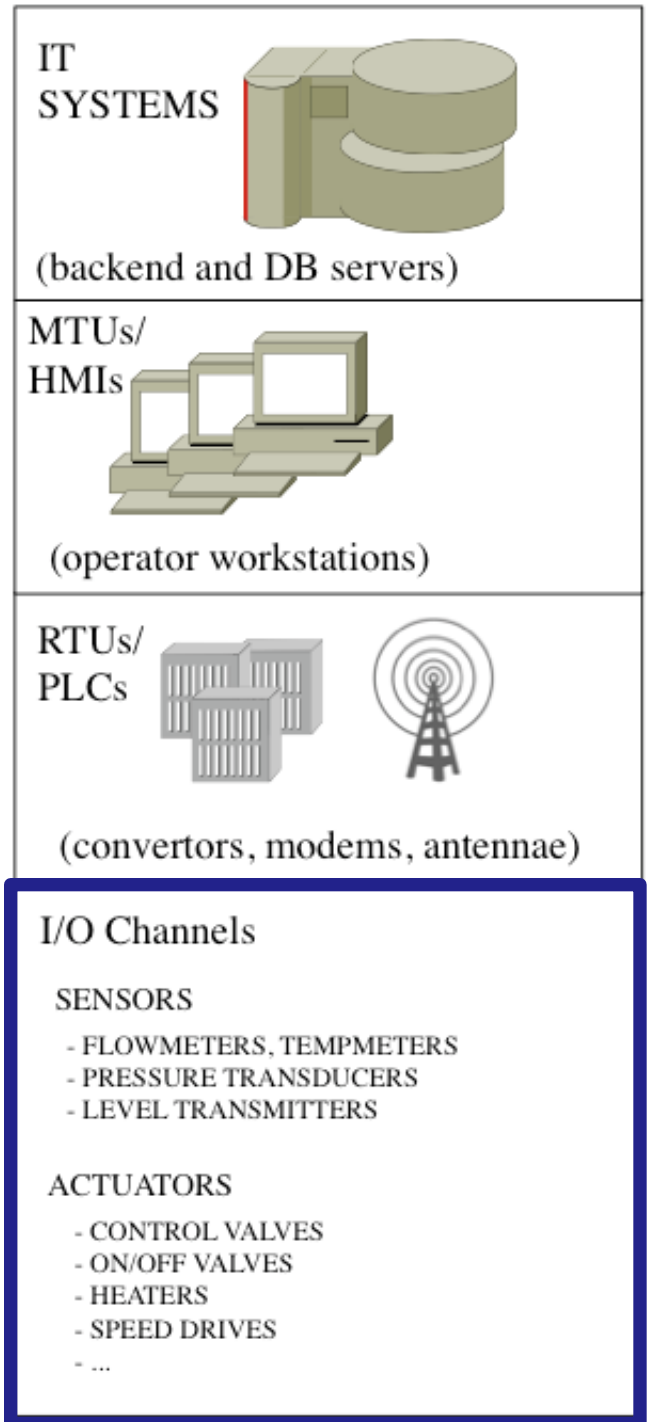
Proposed Methodology

- Foster new theories,
- simulate/emulate case scenarios,
- & validate results using real environments.

Typical SCADA Elements (1/3)

- **Probes/Sensors**: monitoring devices to retrieve measurements related to specific physical phenomena
- **Effectors/Actuators**: control devices, in charge of managing some external devices



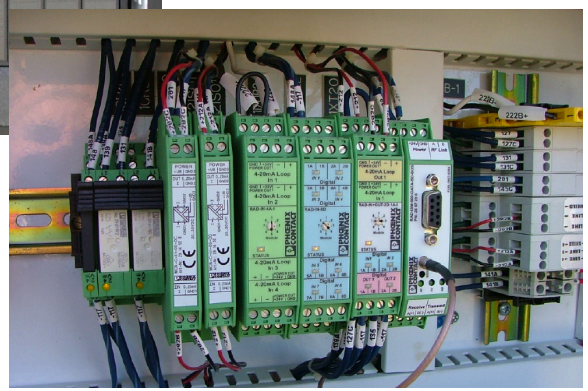


Typical SCADA Elements (2/3)

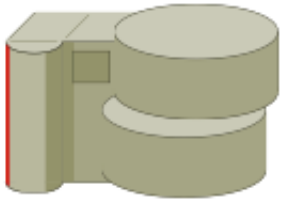
Middleware based on

- **Remote Terminal Units**
- **Programmable Logic Controllers**

to control those devices monitoring/controlling end-points, often deployed far away from the backend

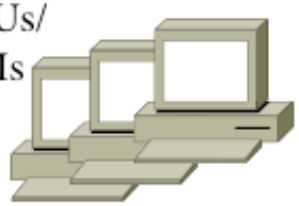


IT SYSTEMS



(backend and DB servers)

MTUs/ HMIs



(operator workstations)

RTUs/ PLCs

(convertors, modems, antennae)

I/O Channels

SENSORS

- FLOWMETERS, TEMPMETERS
- PRESSURE TRANSDUCERS
- LEVEL TRANSMITTERS

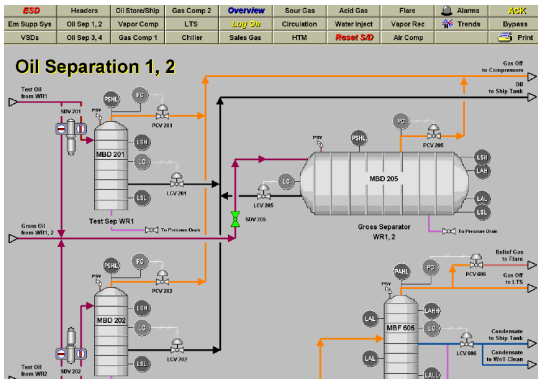
ACTUATORS

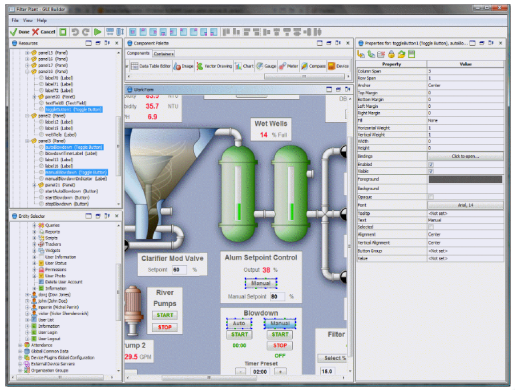
- CONTROL VALVES
- ON/OFF VALVES
- HEATERS
- SPEED DRIVES
- ...

Typical SCADA Elements (3/3)

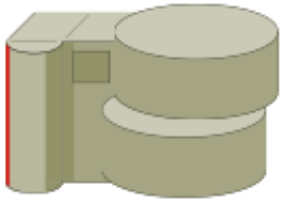
IT/Master Terminal Units/Human Machine Interfaces

- Located at the control center of the organization
- Give access to the management of communications, collection of data, data storage, and control of sensors and actuators via the RTUs/PLCs



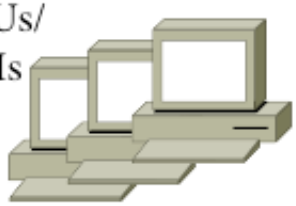


IT SYSTEMS



(backend and DB servers)

MTUs/ HMIs



(operator workstations)

RTUs/ PLCs

(convertors, modems, antennae)

I/O Channels

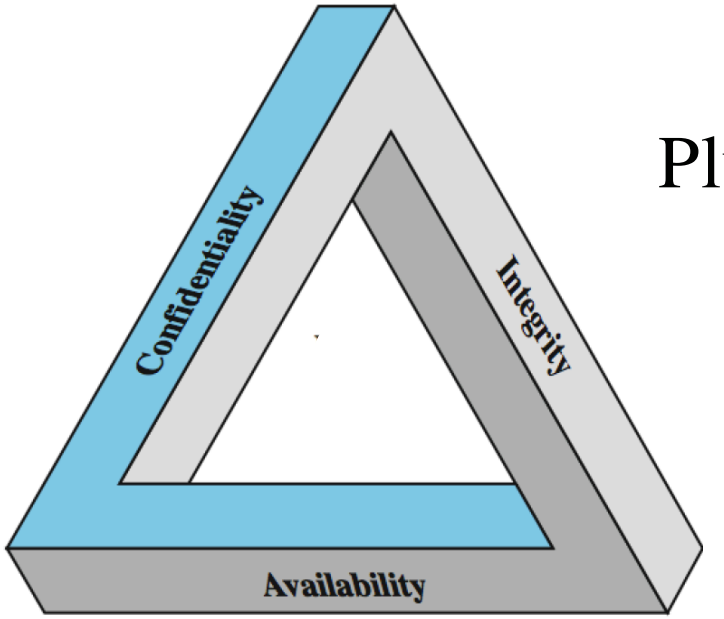
SENSORS

- FLOWMETERS, TEMPMETERS
- PRESSURE TRANSDUCERS
- LEVEL TRANSMITTERS

ACTUATORS

- CONTROL VALVES
- ON/OFF VALVES
- HEATERS
- SPEED DRIVES
- ...

Security Challenges*

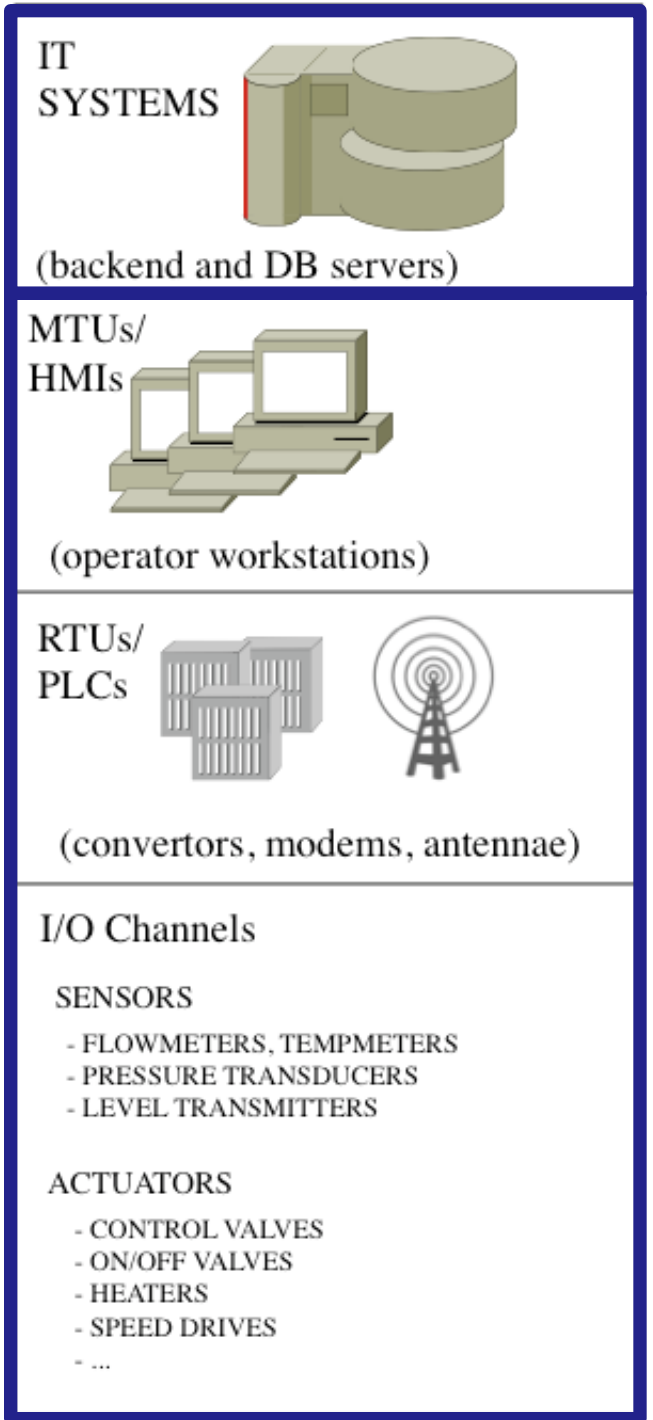


Plus

- Reliability,
- Safety,
- Performance, ...

Asset to protect: Information Process

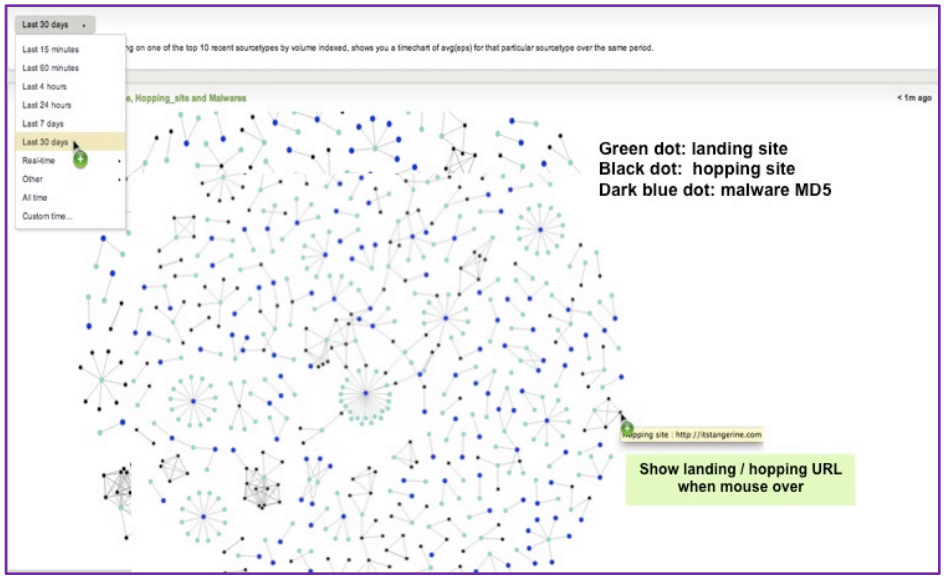
Priority	IT Systems	MTUs to I/O
#1	Confidentiality	Availability
#2	Integrity	Integrity
#3	Availability	Confidentiality



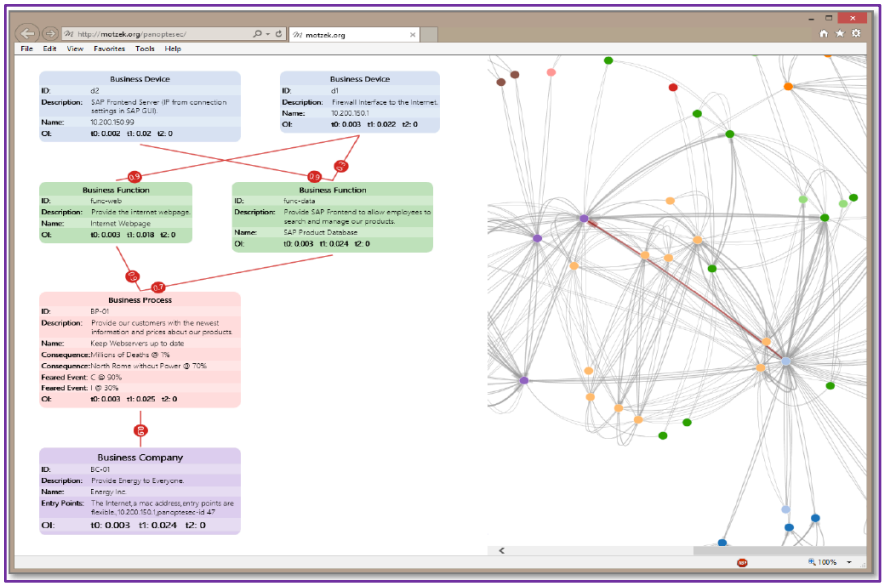
* HIRSCHMANN, Why is Cyber Security Still a Problem? *TOFINO Security Series*

The PANOPTESEC Approach

- Dynamic Risk Assessment
  - Preempt Exploitation of Vulnerabilities
  - Use of Attack & Operational ("*Mission*") Graphs



IT Security Oriented

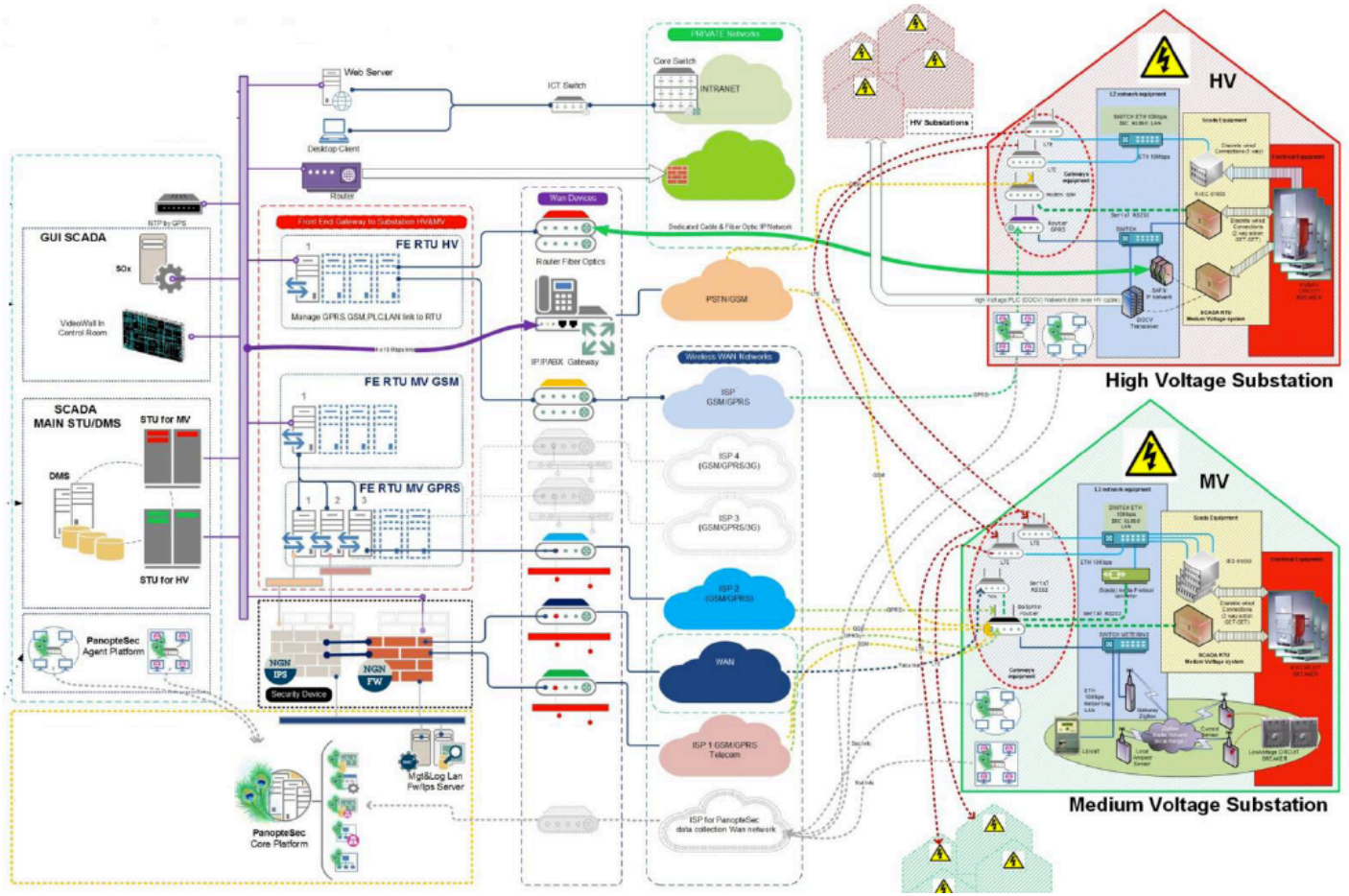


Operational Security Oriented

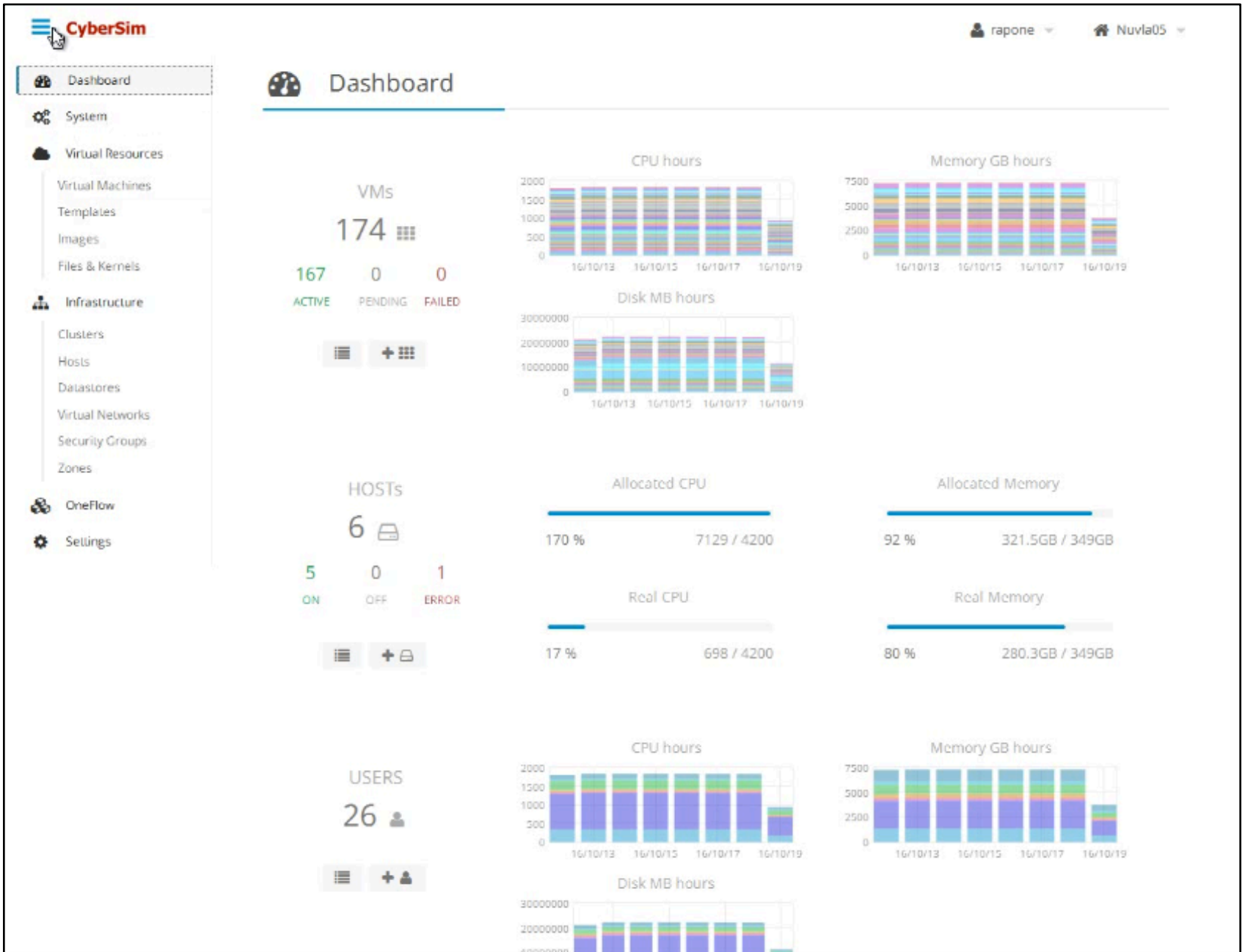
PANOPTESEC

Project Emulation Environment (1/10)

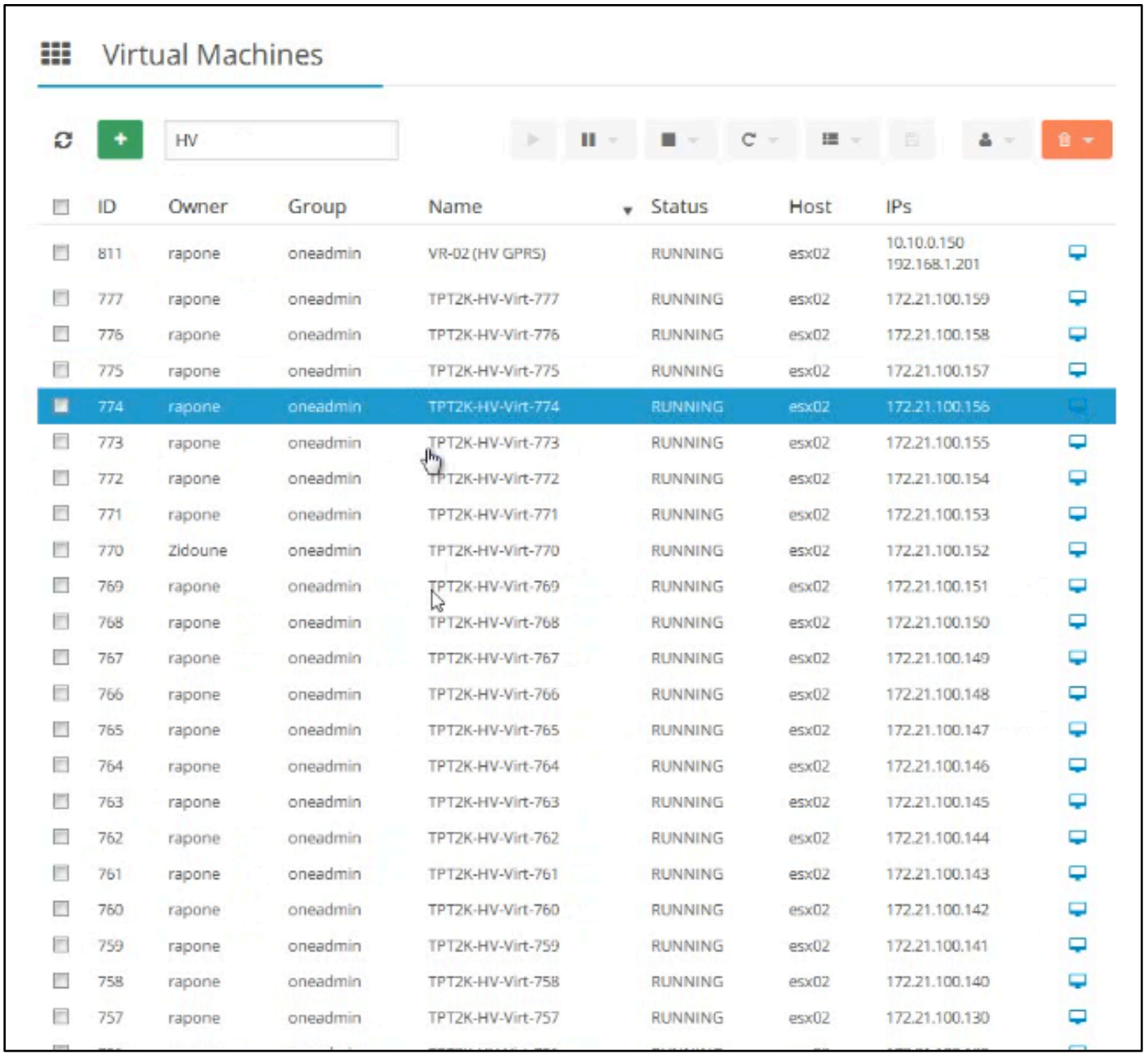
- Real control systems in cold standby mode (Disaster Recovery site)
- Real high/medium voltage substations in test mode
- Cloned SCADA systems for scalability emulation
- Simulated ICT domain for diversity of application experiments (e.g., non-SCADA environments)
- Installed network and security monitoring systems support data collection



Project Emulation Environment (2/10)



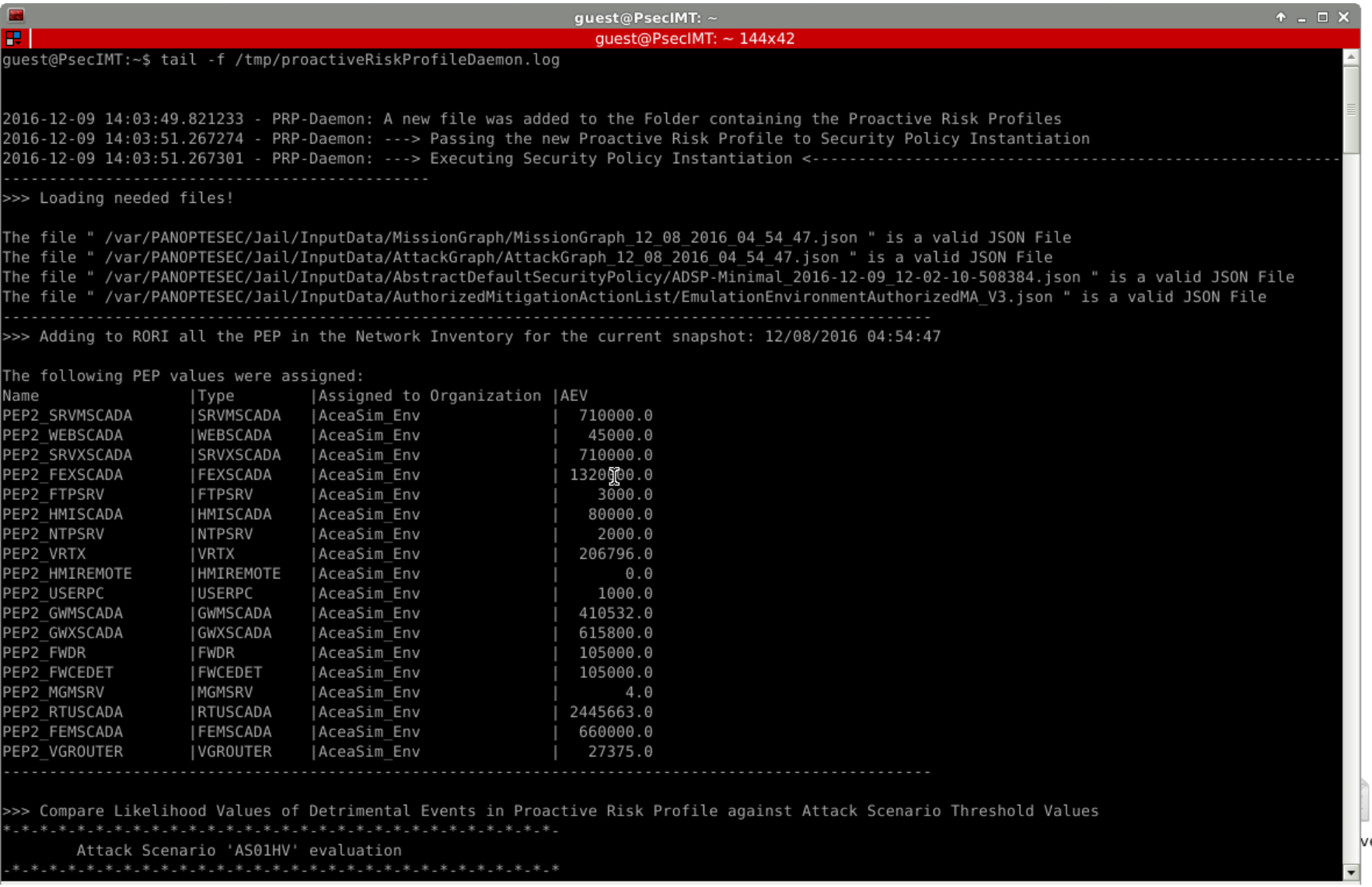
Project Emulation Environment (3/10)



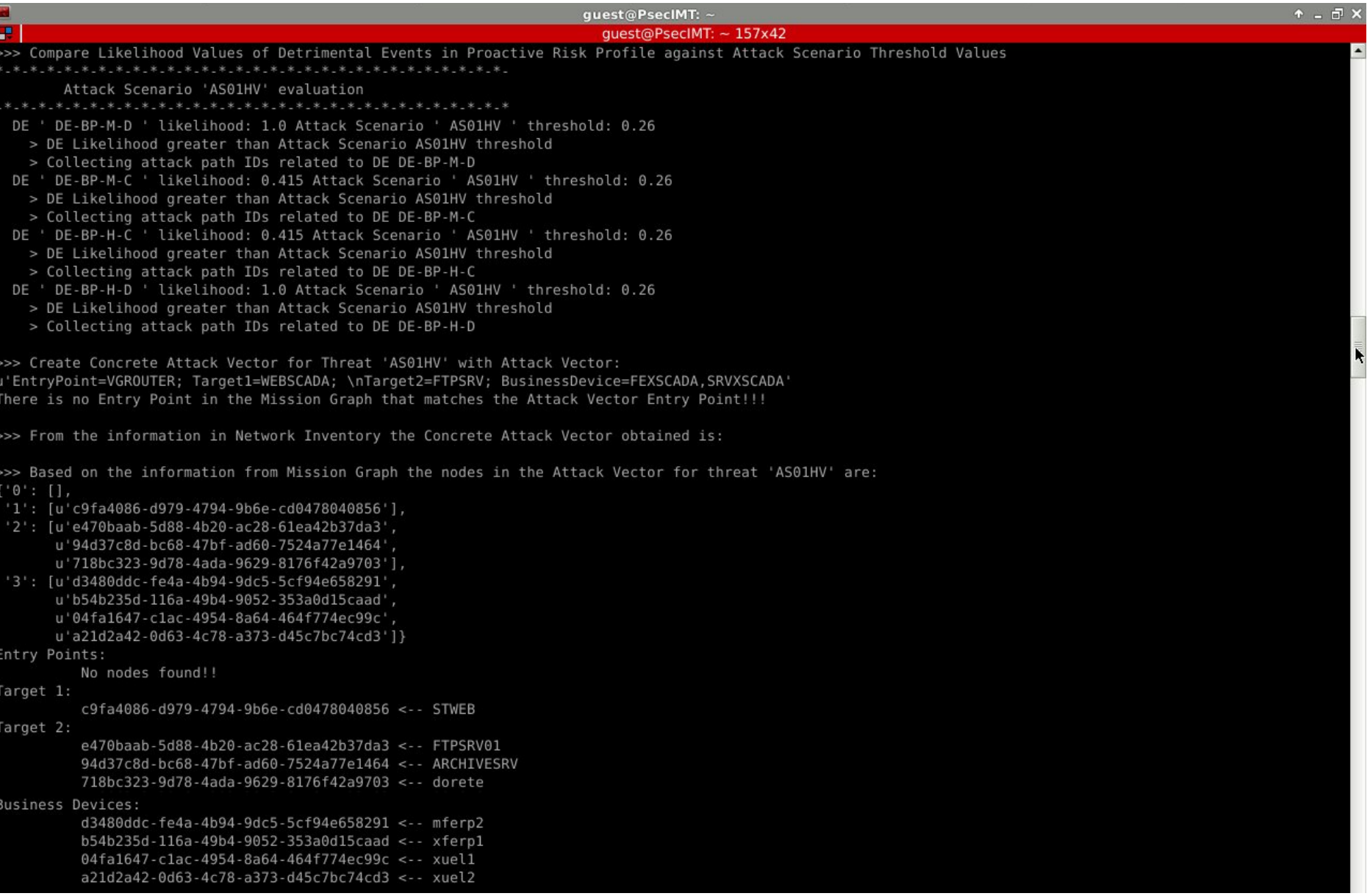
Project Emulation Environment (4/10)

```
14:32:18.807376 ARP, Request who-has 172.21.100.197 tell 172.21.100.10, length 46
14:32:19.547769 ARP, Request who-has 172.21.100.232 tell 172.21.100.10, length 46
14:32:19.547809 ARP, Request who-has 172.21.100.209 tell 172.21.100.10, length 46
14:32:19.553328 ARP, Request who-has 172.21.100.169 tell 172.21.100.10, length 46
14:32:19.557392 ARP, Request who-has 172.21.100.236 tell 172.21.100.10, length 46
14:32:20.562219 ARP, Request who-has 172.21.100.199 tell 172.21.100.10, length 46
14:32:20.563395 ARP, Request who-has 172.21.100.237 tell 172.21.100.10, length 46
14:32:20.572426 ARP, Request who-has 172.21.100.238 tell 172.21.100.10, length 46
14:32:20.874248 IP 192.168.1.4 > 172.21.100.139: ICMP echo request, id 10512, seq 1, length 64
14:32:20.874290 IP 172.21.100.139 > 192.168.1.4: ICMP echo reply, id 10512, seq 1, length 64
14:32:20.876180 IP 192.168.1.4.60996 > 172.21.100.139.sunrpc: Flags [S], seq 2907222934, win 14600,
options [mss 1460,sackOK,TS val 2761366341 ecr 0,nop,wscale 5], length 0
14:32:20.876238 IP 172.21.100.139.sunrpc > 192.168.1.4.60996: Flags [S.], seq 719133229, ack 2907222
935, win 28960, options [mss 1460,sackOK,TS val 488939676 ecr 2761366341,nop,wscale 6], length 0
14:32:20.877093 IP 192.168.1.4.60996 > 172.21.100.139.sunrpc: Flags [.], ack 1, win 457, options [no
p,nop,TS val 2761366341 ecr 488939676], length 0
14:32:20.877154 IP 192.168.1.4.60996 > 172.21.100.139.sunrpc: Flags [P.], seq 1:61, ack 1, win 457,
options [nop,nop,TS val 2761366341 ecr 488939676], length 60
14:32:20.877170 IP 172.21.100.139.sunrpc > 192.168.1.4.60996: Flags [.], ack 61, win 453, options [n
op,nop,TS val 488939677 ecr 2761366341], length 0
14:32:20.877833 IP 172.21.100.139.sunrpc > 192.168.1.4.60996: Flags [P.], seq 1:33, ack 61, win 453,
 options [nop,nop,TS val 488939677 ecr 2761366341], length 32
14:32:20.878647 IP 192.168.1.4.60996 > 172.21.100.139.sunrpc: Flags [.], ack 33, win 457, options [n
op,nop,TS val 2761366342 ecr 488939677], length 0
14:32:20.878664 IP 192.168.1.4.60996 > 172.21.100.139.sunrpc: Flags [F.], seq 61, ack 33, win 457, o
ptions [nop,nop,TS val 2761366342 ecr 488939677], length 0
14:32:20.878781 IP 172.21.100.139.sunrpc > 192.168.1.4.60996: Flags [F.], seq 33, ack 62, win 453, o
ptions [nop,nop,TS val 488939677 ecr 2761366342], length 0
14:32:20.888147 IP 192.168.1.4.60996 > 172.21.100.139.sunrpc: Flags [.], ack 34, win 457, options [n
op,nop,TS val 2761366343 ecr 488939677], length 0
14:32:21.553720 ARP, Request who-has 172.21.100.180 tell 172.21.100.10, length 46
14:32:21.554540 ARP, Request who-has 172.21.100.230 tell 172.21.100.10, length 46
14:32:21.554553 ARP, Request who-has 172.21.100.229 tell 172.21.100.10, length 46
14:32:21.554557 ARP, Request who-has 172.21.100.177 tell 172.21.100.10, length 46
14:32:22.547715 ARP, Request who-has 172.21.100.170 tell 172.21.100.10, length 46
14:32:23.534770 ARP, Request who-has 172.21.100.231 tell 172.21.100.10, length 46
_
```

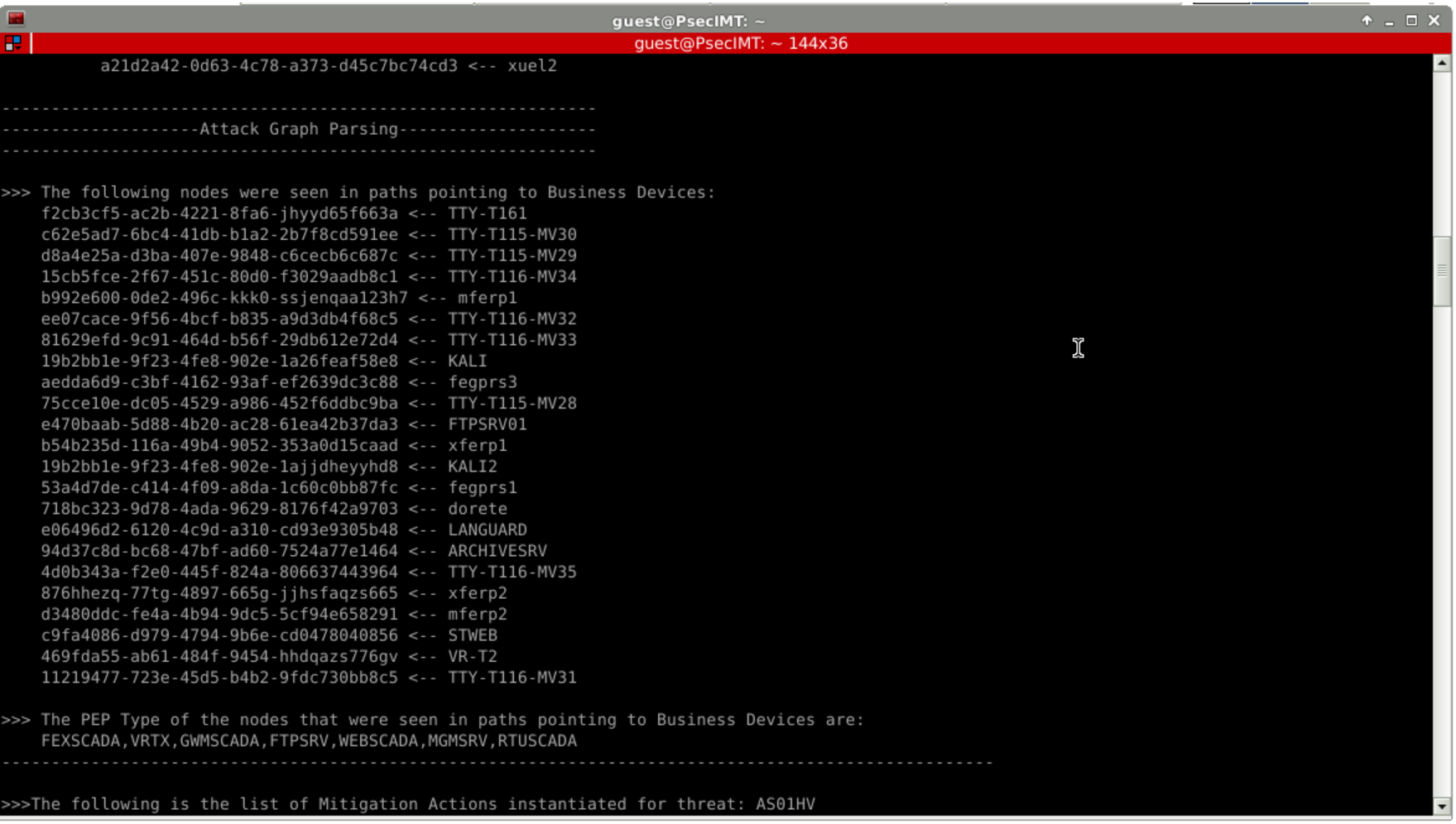
Project Emulation Environment (5/10)



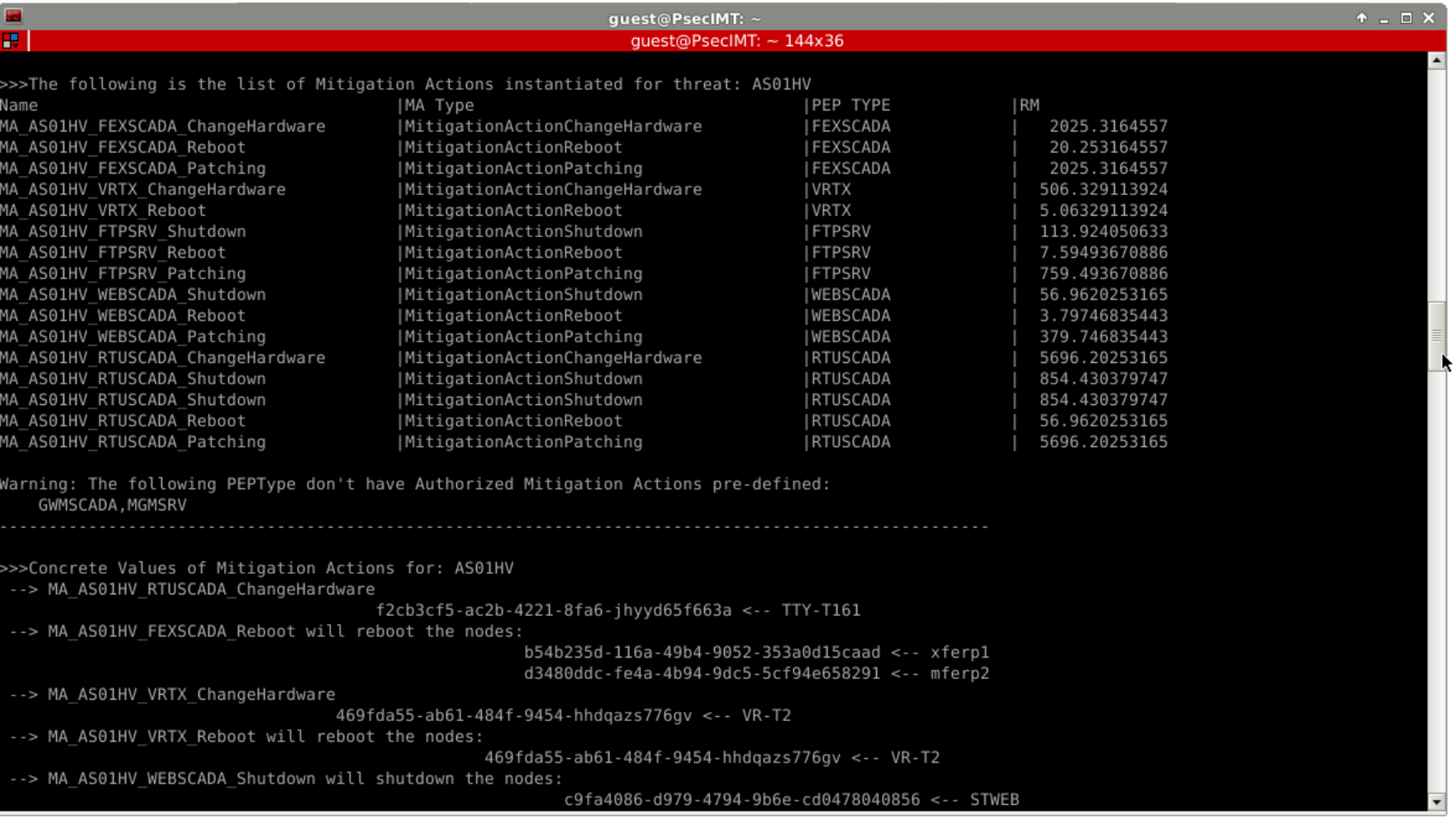
Project Emulation Environment (6/10)



Project Emulation Environment (7/10)



Project Emulation Environment (8/10)



Project Emulation Environment (9/10)

guest@PsecIMT: ~

guest@PsecIMT: ~ 157x42

```
>>>Concrete Values of Mitigation Actions for: AS01HV
 --> MA_AS01HV_RTUSCADA_ChangeHardware
                                  f2cb3cf5-ac2b-4221-8fa6-jhyyd65f663a <-- TTY-T161
 --> MA_AS01HV_FEXSCADA_Reboot will reboot the nodes:
                                              b54b235d-116a-49b4-9052-353a0d15caad <-- xferp1
                                              d3480ddc-fe4a-4b94-9dc5-5cf94e658291 <-- mferp2
 --> MA_AS01HV_VRTX_ChangeHardware
                               469fda55-ab61-484f-9454-hhdqazs776gv <-- VR-T2
 --> MA_AS01HV_VRTX_Reboot will reboot the nodes:
                                            469fda55-ab61-484f-9454-hhdqazs776gv <-- VR-T2
 --> MA_AS01HV_WEBSCADA_Shutdown will shutdown the nodes:
                                                c9fa4086-d979-4794-9b6e-cd0478040856 <-- STWEB
 --> MA_AS01HV_FTPSRV_Patching will patch the nodes:
                                           718bc323-9d78-4ada-9629-8176f42a9703 <-- dorete, against CVE:CVE-2008-4250
                                           718bc323-9d78-4ada-9629-8176f42a9703 <-- dorete, against CVE:CVE-2006-3439
                                           e470baab-5d88-4b20-ac28-61ea42b37da3 <-- FTPSRV01, against CVE:CVE-2004-2687
                                           94d37c8d-bc68-47bf-ad60-7524a77e1464 <-- ARCHIVESRV, against CVE:CVE-2004-2687
                                           e470baab-5d88-4b20-ac28-61ea42b37da3 <-- FTPSRV01, against CVE:CVE-2010-2075
                                           94d37c8d-bc68-47bf-ad60-7524a77e1464 <-- ARCHIVESRV, against CVE:CVE-2010-2075
 --> MA_AS01HV_RTUSCADA_Shutdown will shutdown the nodes:
                                               c62e5ad7-6bc4-41db-b1a2-2b7f8cd591ee <-- TTY-T115-MV30
                                               d8a4e25a-d3ba-407e-9848-c6cecb6c687c <-- TTY-T115-MV29
                                               15cb5fce-2f67-451c-80d0-f3029aadb8c1 <-- TTY-T116-MV34
                                               ee07cace-9f56-4bcf-b835-a9d3db4f68c5 <-- TTY-T116-MV32
                                               81629efd-9c91-464d-b56f-29db612e72d4 <-- TTY-T116-MV33
                                               75cce10e-dc05-4529-a986-452f6ddbc9ba <-- TTY-T115-MV28
                                               4d0b343a-f2e0-445f-824a-806637443964 <-- TTY-T116-MV35
                                               11219477-723e-45d5-b4b2-9fdc730bb8c5 <-- TTY-T116-MV31
 --> MA_AS01HV_WEBSCADA_Patching will patch the nodes:
                                             c9fa4086-d979-4794-9b6e-cd0478040856 <-- STWEB, against CVE:CVE-2008-4250
                                             c9fa4086-d979-4794-9b6e-cd0478040856 <-- STWEB, against CVE:CVE-2006-3439
 --> MA_AS01HV_RTUSCADA_Patching will patch the nodes:
                                             c62e5ad7-6bc4-41db-b1a2-2b7f8cd591ee <-- TTY-T115-MV30, against CVE:CVE-2004-0185
                                             d8a4e25a-d3ba-407e-9848-c6cecb6c687c <-- TTY-T115-MV29, against CVE:CVE-2004-0185
                                             15cb5fce-2f67-451c-80d0-f3029aadb8c1 <-- TTY-T116-MV34, against CVE:CVE-2004-0185
                                             ee07cace-9f56-4bcf-b835-a9d3db4f68c5 <-- TTY-T116-MV32, against CVE:CVE-2004-0185
                                             81629efd-9c91-464d-b56f-29db612e72d4 <-- TTY-T116-MV33, against CVE:CVE-2004-0185
                                             75cce10e-dc05-4529-a986-452f6ddbc9ba <-- TTY-T115-MV28, against CVE:CVE-2004-0185
                                             4d0b343a-f2e0-445f-824a-806637443964 <-- TTY-T116-MV35, against CVE:CVE-2004-0185
                                             11219477-723e-45d5-b4b2-9fdc730bb8c5 <-- TTY-T116-MV31, against CVE:CVE-2004-0185
                                             c62e5ad7-6bc4-41db-b1a2-2b7f8cd591ee <-- TTY-T115-MV30, against CVE:CVE-2003-1327
                                             d8a4e25a-d3ba-407e-9848-c6cecb6c687c <-- TTY-T115-MV29, against CVE:CVE-2003-1327
```

Project Emulation Environment (10/10)

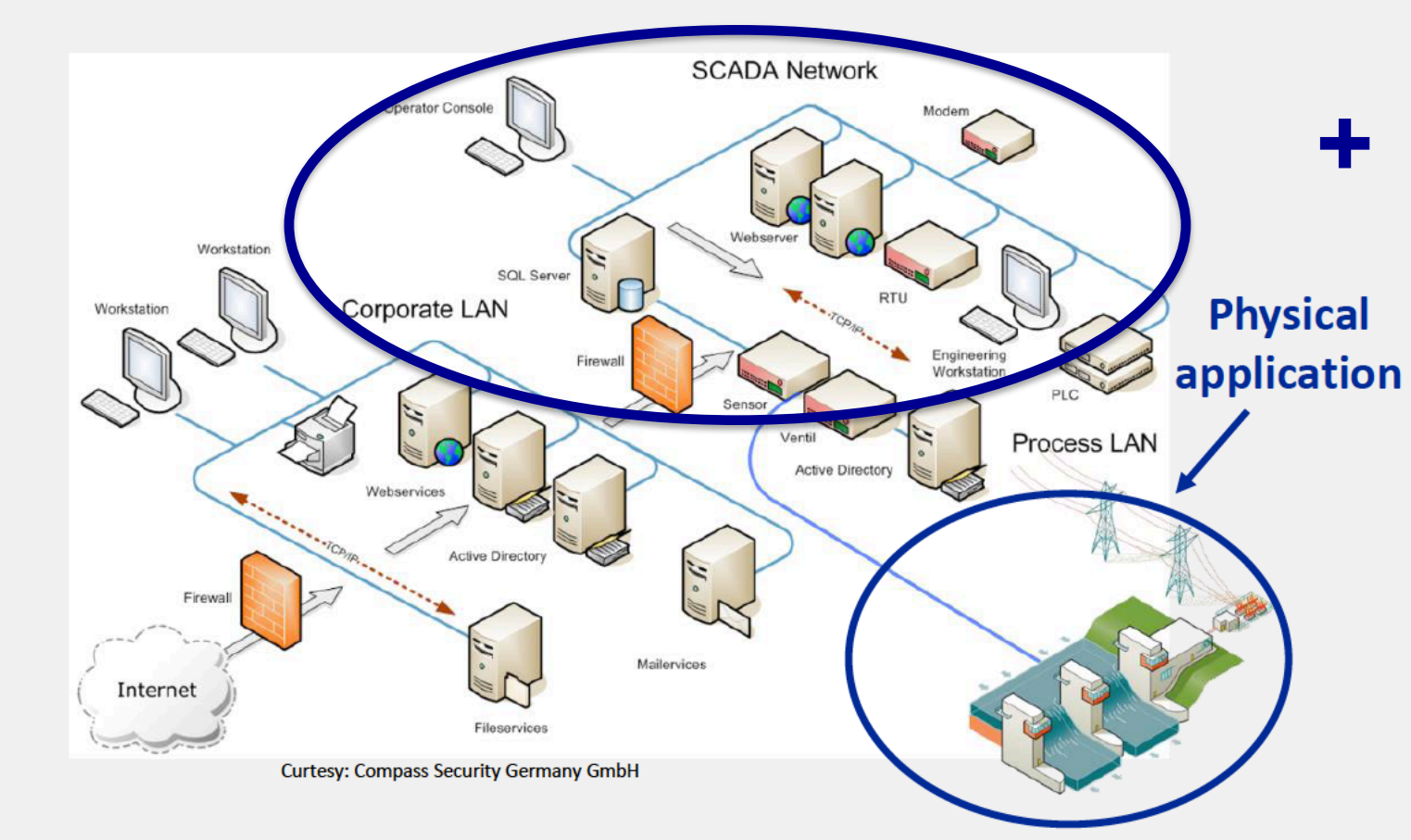
guest@PsecIMT: ~

guest@PsecIMT: ~ 144x41

```
Checking Parameters Given
Performing RORI evaluation for:
'AS01HV' incident at the organization: 'AceaSim_Env'
-------------------------------------------------------------------------------------------------
Individual Results
ID  |NAME                                        |EQUIPMENT ID|RISK MITIGATION|RORI
MA_AS01HV_WEBSCADA_Shutdown|MA_AS01HV_WEBSCADA_Shutdown                  |WEBSCADA  |  0.0057| 1.52942
MA_AS01HV_WEBSCADA_Reboot|MA_AS01HV_WEBSCADA_Reboot                    |WEBSCADA  |  0.0004|0.106079
MA_AS01HV_WEBSCADA_Patching|MA_AS01HV_WEBSCADA_Patching                |WEBSCADA  |  0.0380|10.160842
MA_AS01HV_FEXSCADA_ChangeHardware|MA_AS01HV_FEXSCADA_ChangeHardware          |FEXSCADA  |  0.2025|47.832309
MA_AS01HV_FEXSCADA_Reboot|MA_AS01HV_FEXSCADA_Reboot                    |FEXSCADA  |  0.0020|0.510124
MA_AS01HV_FEXSCADA_Patching|MA_AS01HV_FEXSCADA_Patching                |FEXSCADA  |  0.2025|54.258792
MA_AS01HV_RTUSCADA_ChangeHardware|MA_AS01HV_RTUSCADA_ChangeHardware          |RTUSCADA  |  0.5696|152.224635
MA_AS01HV_RTUSCADA_Reboot|MA_AS01HV_RTUSCADA_Reboot                    |RTUSCADA  |  0.0057| 1.52942
MA_AS01HV_FTPSRV_Shutdown|MA_AS01HV_FTPSRV_Shutdown                    |FTPSRV    |  0.0114|3.060182
MA_AS01HV_FTPSRV_Reboot|MA_AS01HV_FTPSRV_Reboot                      |FTPSRV    |  0.0008|0.213501
MA_AS01HV_FTPSRV_Patching|MA_AS01HV_FTPSRV_Patching                    |FTPSRV    |  0.0759|20.335106
MA_AS01HV_RTUSCADA_Shutdown|MA_AS01HV_RTUSCADA_Shutdown                  |RTUSCADA  |  0.0854|22.933236
MA_AS01HV_VRTX_ChangeHardware|MA_AS01HV_VRTX_ChangeHardware              |VRTX      |  0.0506|13.065779
MA_AS01HV_VRTX_Reboot|MA_AS01HV_VRTX_Reboot                        |VRTX      |  0.0005|0.107395
MA_AS01HV_RTUSCADA_Patching|MA_AS01HV_RTUSCADA_Patching                  |RTUSCADA  |  0.5696|152.868995 <-- Best RORI index
-------------------------------------------------------------------------------------------------
The selected combination criteria is to only combine the Mitigation Actions which RORI is over the RORI average
The individual RORI average is:',rori_avg
The following Mitigation Actions are over such average:
- MA_AS01HV_FEXSCADA_ChangeHardware - MA_AS01HV_FEXSCADA_Patching - MA_AS01HV_RTUSCADA_ChangeHardware - MA_AS01HV_RTUSCADA_Patching -
-------------------------------------------------------------------------------------------------
Combined Results
Mitigation Actions            |RM      |RORI Index
[MA_AS01HV_RTUSCADA_Patching
MA_AS01HV_FEXSCADA_Patching]|0.6708  |179.8242 <-- Best RORI index
[MA_AS01HV_FEXSCADA_Patching
MA_AS01HV_RTUSCADA_ChangeHardware]|0.6708  |179.1118
[MA_AS01HV_RTUSCADA_Patching
MA_AS01HV_FEXSCADA_ChangeHardware]|0.6708  |168.1711
[MA_AS01HV_FEXSCADA_ChangeHardware
MA_AS01HV_RTUSCADA_ChangeHardware]|0.6708  |167.5167
-------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------
19 Response Plans were generated for threat:AS01HV
Response Plans Saved in: /var/PANOPTESEC/RPGenerator/ResponsePlans/12-08-2016_04-54-47/AS01HV01HV_FTPSRV_Shutdown will shutdown the nodes:
```

Evolution

- Protect, as well, from threats that are affecting physical sensors and actuators
  - *In other words …*

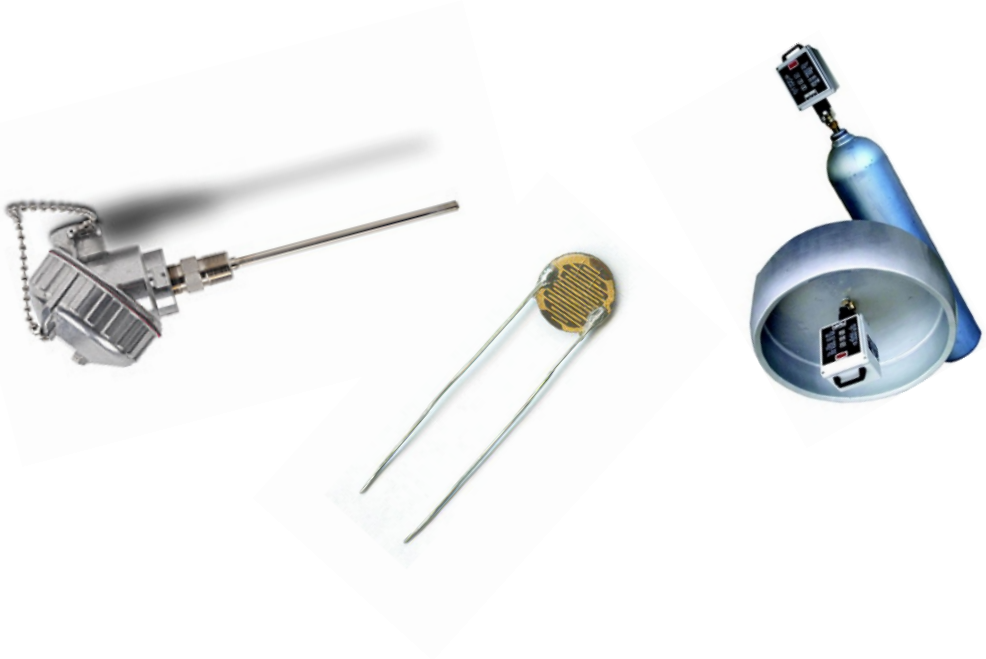


Source: *Hacking Chemical Plants for Competition and Extortion*, Krotofil and Larsen, DefCon23, 2015.

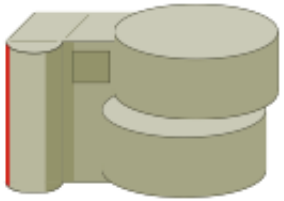
Physical Elements

- **Probes/Sensors**: monitoring devices in to retrieve measurements related to specific physical phenomena

- **Effectors/Actuators**: control devices, in charge of managing some external devices

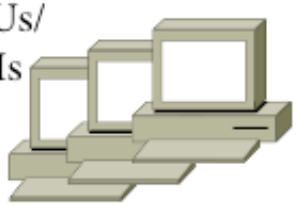


IT SYSTEMS



(backend and DB servers)

MTUs/ HMIs



(operator workstations)

RTUs/ PLCs

(convertors, modems, antennae)

I/O Channels

SENSORS

- FLOWMETERS, TEMPMETERS
- PRESSURE TRANSDUCERS
- LEVEL TRANSMITTERS

ACTUATORS

- CONTROL VALVES
- ON/OFF VALVES
- HEATERS
- SPEED DRIVES
- ...

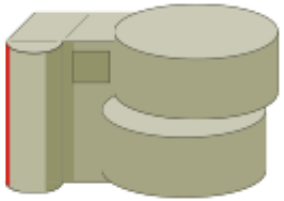
Physical Elements

Middleware based on:

- **Remote Terminal Units**
- **Programmable Logic Controllers**

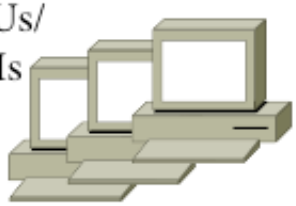
to control a myriad (thousand to million) of devices monitoring/controlling end-points, often deployed far away (hundreds to thousands of km) from the backend

IT SYSTEMS



(backend and DB servers)

MTUs/ HMIs



(operator workstations)

RTUs/ PLCs

(convertors, modems, antennae)

I/O Channels

SENSORS

- FLOWMETERS, TEMPMETERS
- PRESSURE TRANSDUCERS
- LEVEL TRANSMITTERS

ACTUATORS

- CONTROL VALVES
- ON/OFF VALVES
- HEATERS
- SPEED DRIVES
- ...

Outline

- Brief Introduction
- *Cyber-Physical Systems*
- Feedback Control Verification
- Summary & Perspectives

Fundamental Questions …

- What are Cyber-Physical Systems (CPSs)?
- Are CPSs new?
- How CPS security differs from traditional IT security?

What are CPSs?

- Systems that monitor behavior of physical processes and *take actions to correct those behaviors*

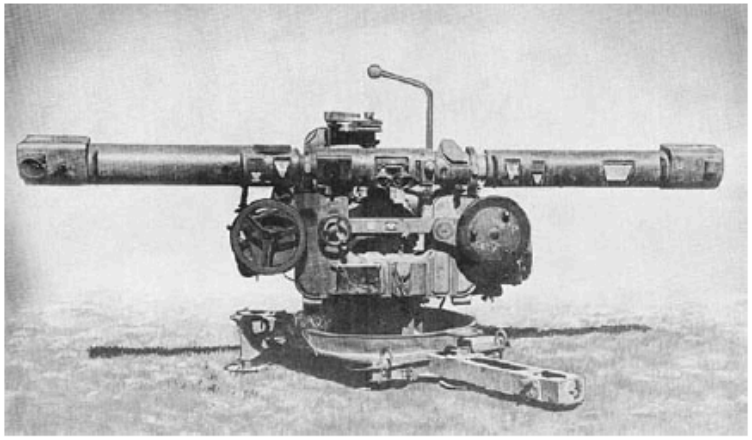
... but also

Another definition

"Systems with integrated computational and physical capabilities that **can interact with humans** through many new modalities"

Are CPSs new?

Short answer: **No, they are not***



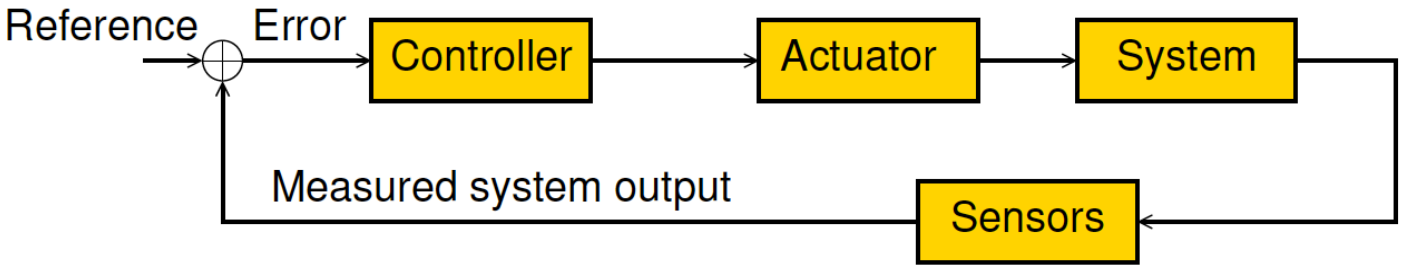
The Kommandogerät 40 (Anti-aircraft gun - World War II)

- Integration of mechanics, electronics, communication
- Human(s) in the loop (required at least five humans to operate)

** Cyber–Physical Systems: A Perspective at the Centennial. Kim and Kumar. Proceedings of the IEEE, Vol. 100, pages 1287-1308, May 2012.*

The key ingredient in a CPS: Control

- **Control** means making a (dynamical) system to work as required
- **Feedback** is used to compute a corrective **control action** based on the distance between a *reference signal* and the *system output*



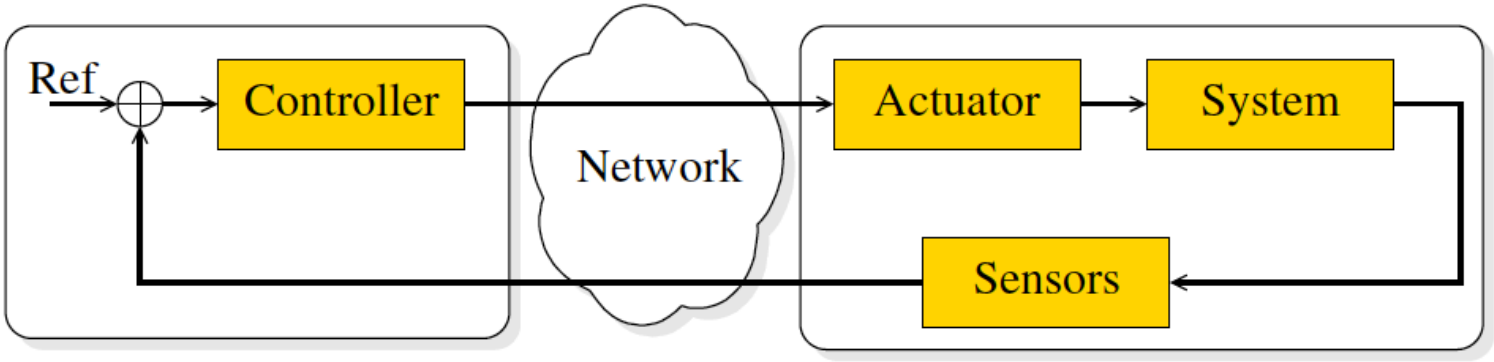
- Examples: dynamically follow a trajectory (robotics), regulate a temperature, regulate the sending rate of a TCP sender (TCP cong. control), controlling a pendulum in its unstable equilibrium, etc.

Networked Control System

- From a methodological standpoint, we can model a CPS using a Networked-Control System (NCS)

NCS definition

Control system whose control loops are connected through a communication network



Traditional Issues Studied in the NCS Literature

- Stabilizing a system under network delays & packet losses
- Techniques to limit data rate (e.g., from control to plant)
- Energy efficient networking for Wireless NCS
- Security?
  - Since the *stuxnet* incident, the control community seems to be heavily working as well on security issues of CPSs

CPS Vulnerabilities

- Traditional Security Issues at the Cyber layer
  - Unencrypted communications
  - Controller settings manually configured (remotely or in person)
  - Default usernames and passwords
  - …

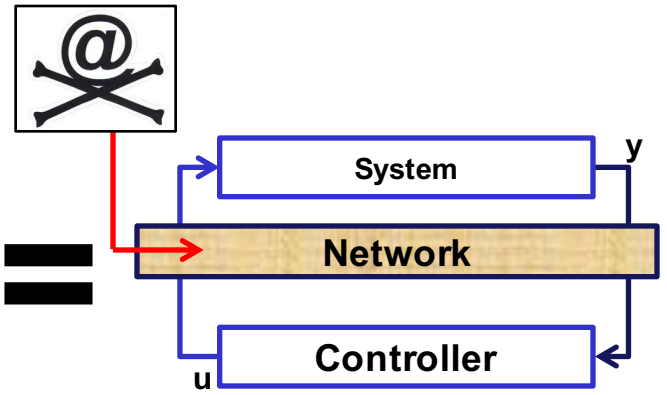
- Attack Surface
  - Physical & control (**Physical**-layer)
  - Communication & network (**Cyber**-layer)
  - Supervisory & management (**Human**-layer)
  - …

- Attack Vectors
  - Data (Control & **Measurements** / Actuators & Sensors)
  - Estimations & **Orders** (Controller & HMIs)
  - ...

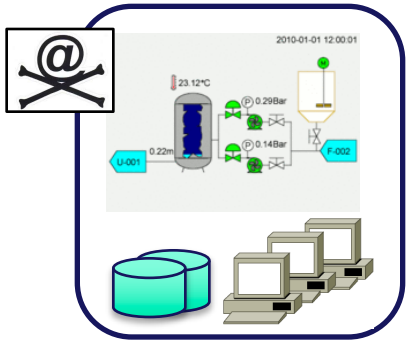
Putting all Together …

People & Control Loops

Information and Communications Technologies (ICT)

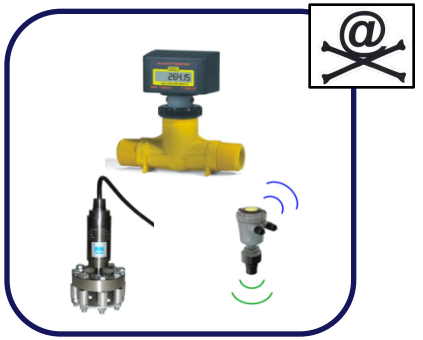


Networked Control System (NCS)



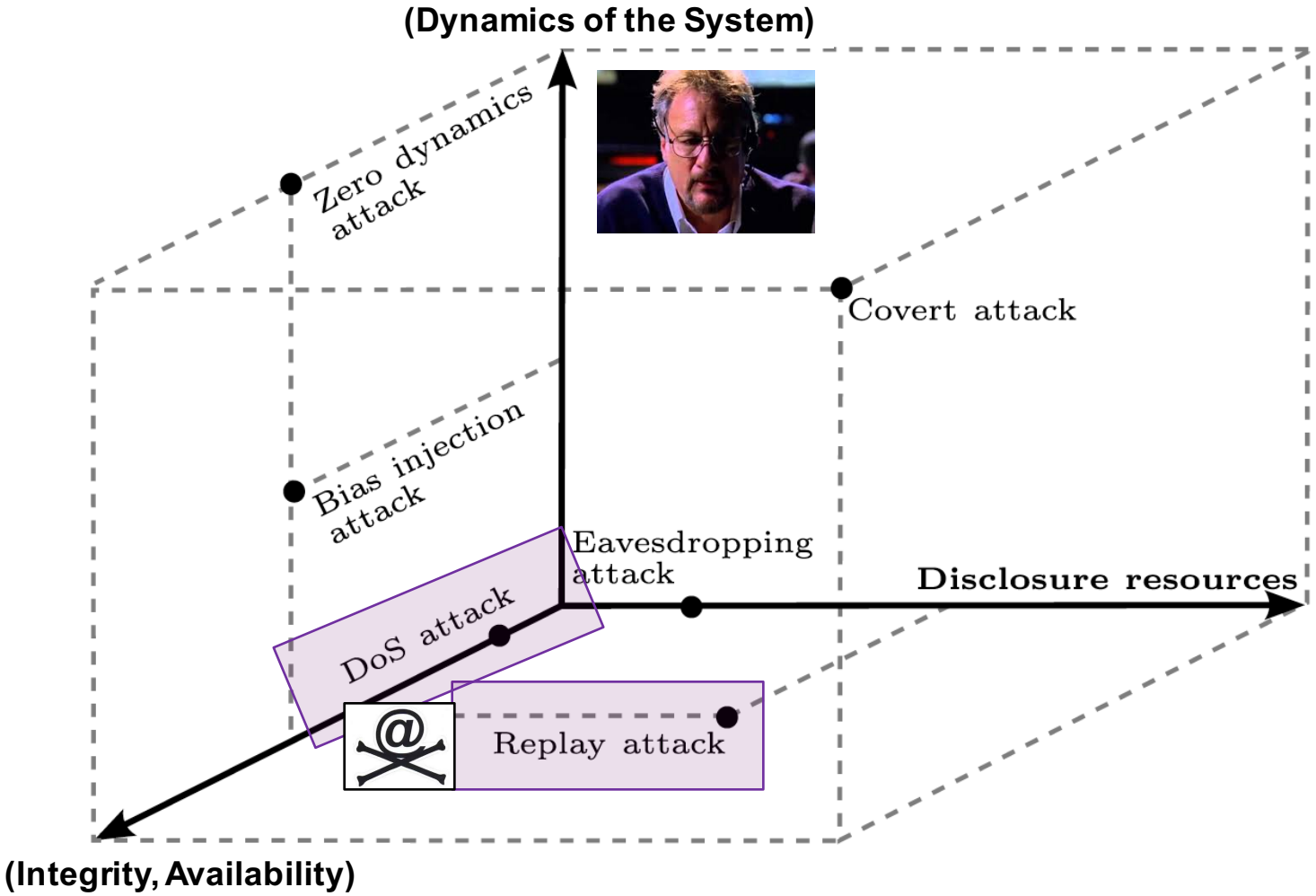
Management Systems

Programmable Automata



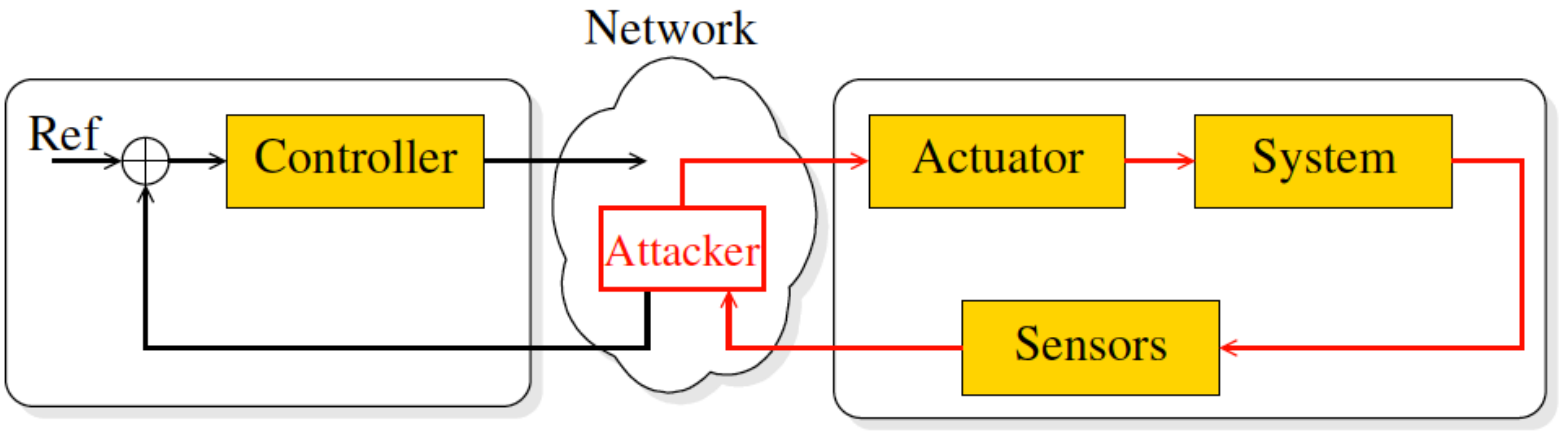
Sensors & Actuators

Sample Attacks*



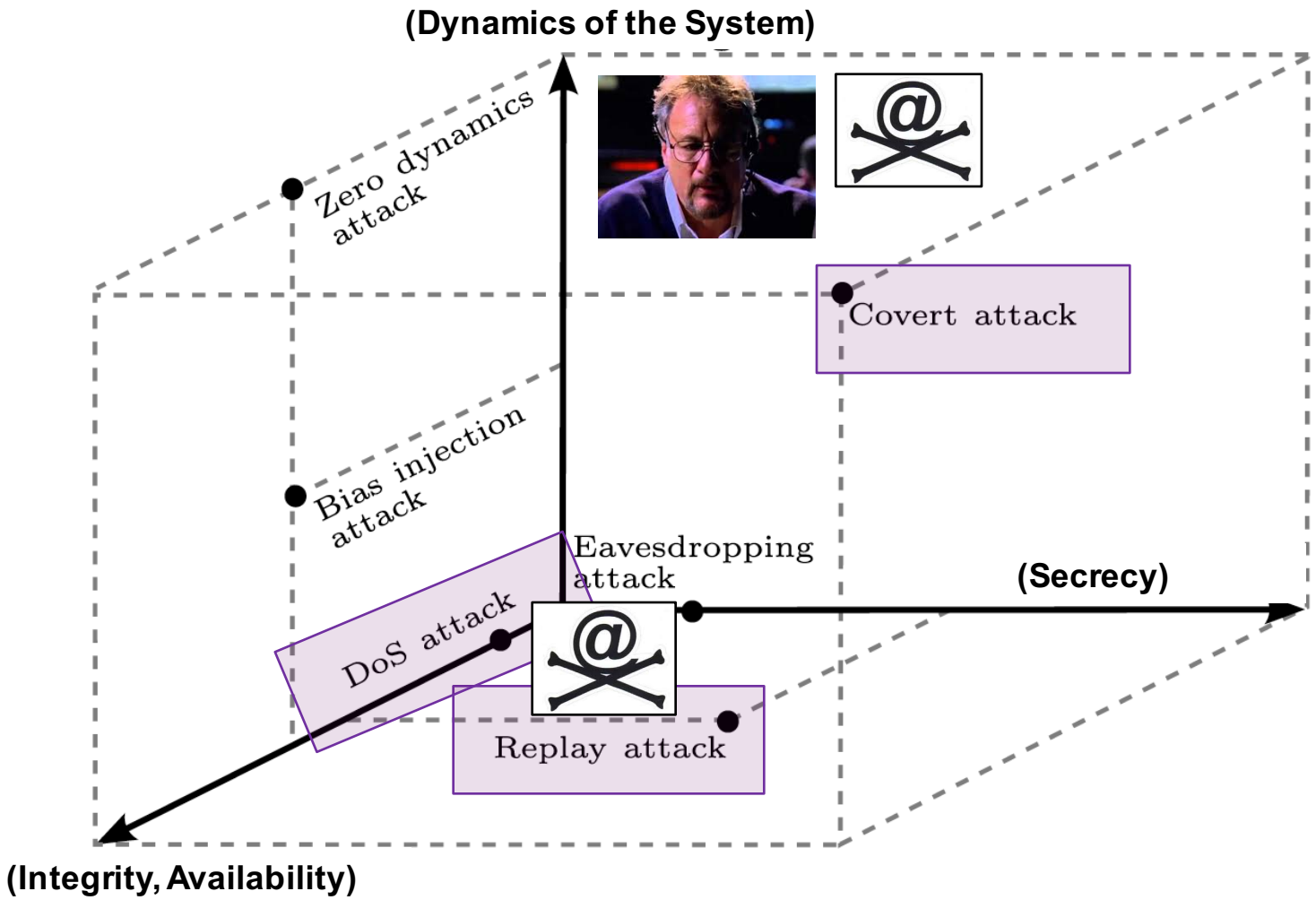
** A secure control framework for resource-limited adversaries. Texeira et al., Automatica, 51(1):135-148, 2015.*

Replay Attack



- Step 1: Sensors output is recorded
- Step 2: Recorded sensors output is replayed and sent to the controller
- Step 3: A control signal is sent to disrupt system functionalities

Sample Attacks*



** A secure control framework for resource-limited adversaries. Texeira et al., Automatica, 51(1):135-148, 2015.*

Prevention of CPS Attacks

- A well-designed control system shall resist external disturbances (failures & attacks), to a certain degree
- Several control-theoretic techniques to prevent cyber-physical attacks have been proposed in the literature*
- Most of the techniques aim at injecting authentication to the control signal & discover anomalous measurements
  - E.g., use a noisy control authentication signal to detect integrity attacks on sensor measurements
  - In the following, we elaborate further on the aforementioned technique

** A survey on the security of cyber-physical systems. Wu, Sun, and Chen. Control Theory and Technology, 14(1):2–10, February 2016.*

Outline

- Brief Introduction
- Cyber-Physical Systems
- ***Feedback Control Verification***
- Summary & Perspectives

Revisiting a Watermark-based Detection Scheme to Handle Cyber-Physical Attacks*

Joint work with

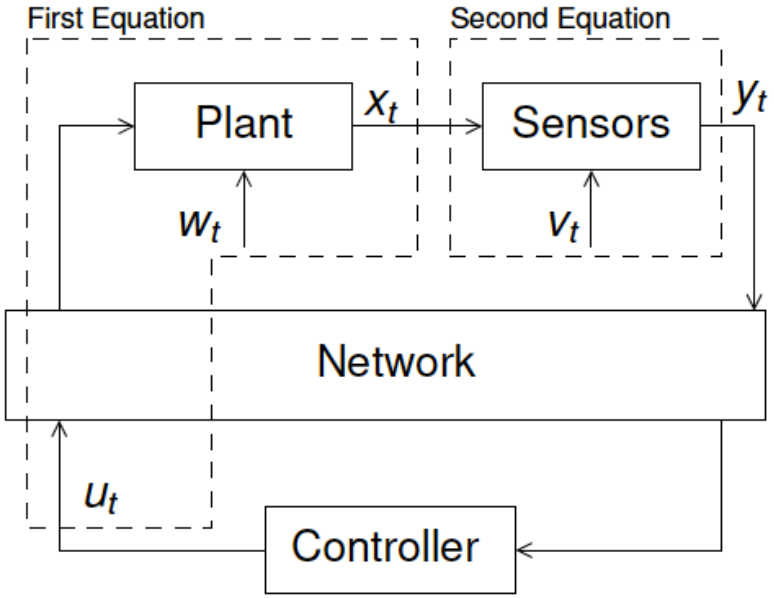
Jose Rubio-Hernan & Luca de Cicco

** 11th International Conference on Availability, Reliability and Security (ARES 2016), August 2016. (Best Paper Runner-Up Award)*

The Mo et al. Approach* (1/2)

System is modeled as follows:

$$x_{t+1} = Ax_t + Bu_t + w_t$$
$$y_t = Cx_t + v_t$$



- x_t is the state vector in $\mathbb{R}^n$
- $u_t \in \mathbb{R}^p$ is the control action
- y_t system output vector in $\mathbb{R}^m$

** Physical Authentication of Control Systems. Mo, Weerakkody and Sinopoli. IEEE Control Systems, Vol. 35, pages 93–109, 2015.*

The Mo et al. Approach* (2/2)

- The control signal input of the plant is:

$$u_t = u_t^* + \Delta u_t$$

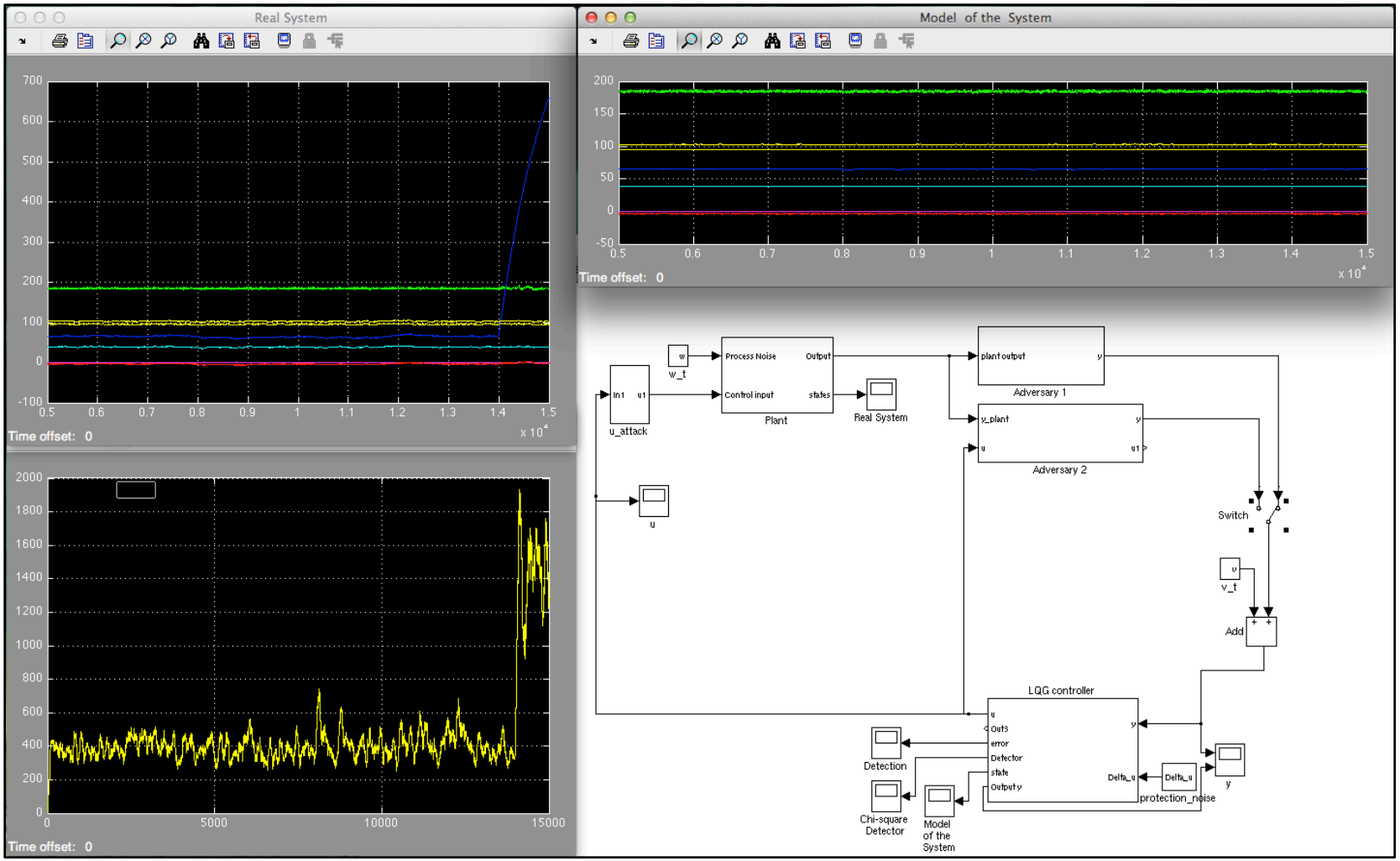
- $\Delta u \in \mathbb{R}^p$ are p Gaussian stationary processes independent from the noises
- A detector is used at the controller that computes alarms based on Kalman filter residues

$$g_t = \sum_{i=t-w+1}^{t} (y_i - C\hat{x}_{i|i-1})^T \mathcal{P}^{-1} (y_i - C\hat{x}_{i|i-1}) \qquad (1)$$

- In normal operation (no attack) g_t is small

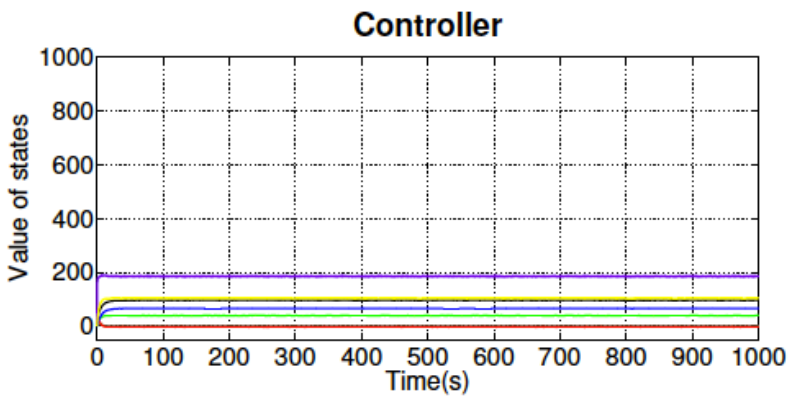
** Physical Authentication of Control Systems. Mo, Weerakkody and Sinopoli. IEEE Control Systems, Vol. 35, pages 93–109, 2015.*

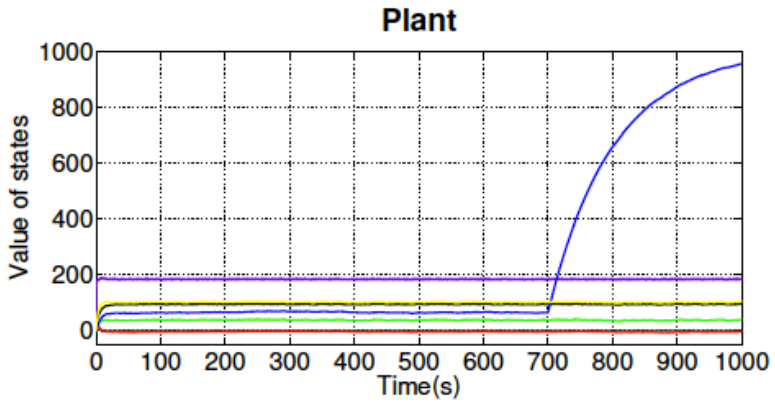
Simulating the Approach in Matlab/Simulink

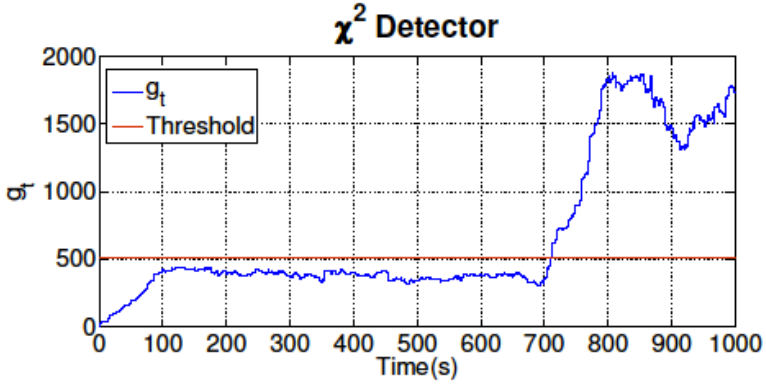


Validating the Approach in Matlab/Simulink

A replay attack is started at time $t = 700s$







- The attacker disrupt system dynamics (the controller does not perceive it)
- The attack is detected

Uncovered Issues

- Via simulation, we can show that a cyber-physical adversary is able to escape the detector

Cyber-physical Adversary

An attacker that is able to eavesdrop the messages containing the output of the controller with the intention of improving its knowledge about the system model

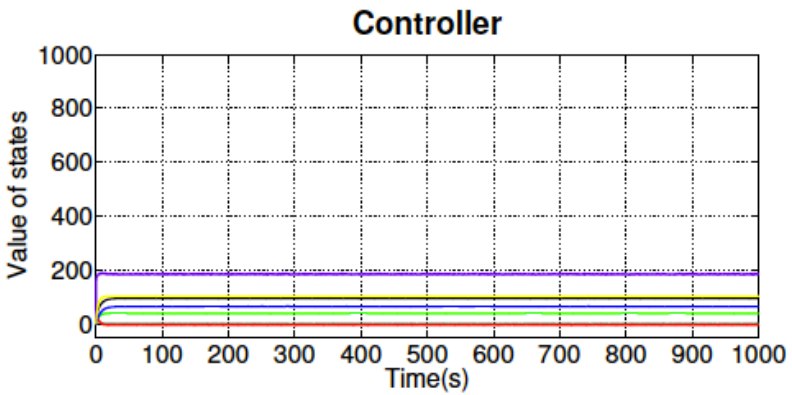
- Such attacker can leverage system identification tools to gather the system model and extract the watermark Δu_t from $u_t = u_t^* + \Delta u_t$
- The extracted watermark can be used to authenticate messages to disrupt system dynamics

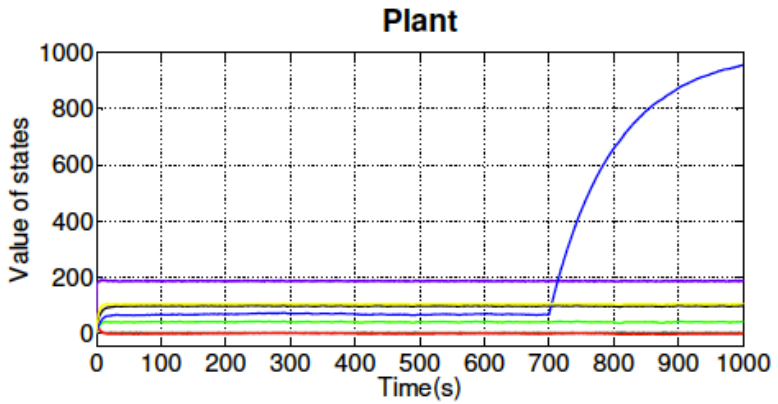
An Implementation of our Proposed Attack

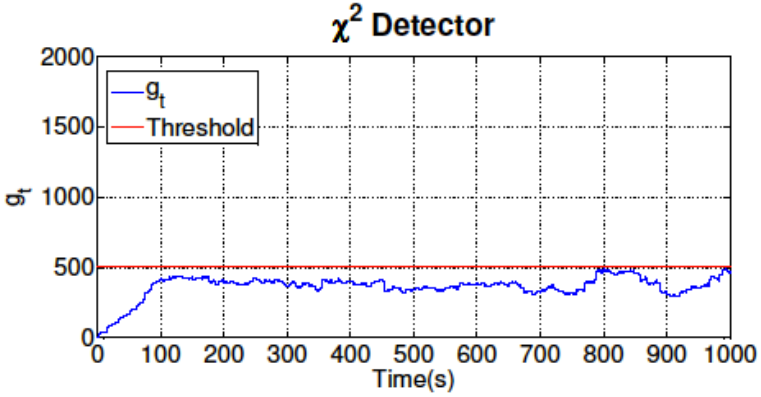
- The attacker eavesdrops u_t and y_t
- A Least Mean Square (LMS) filter is used to get an input-output model $\mathcal{W}$ from u_t to y_t
- By doing so the watermark Δu_t is obtained
- An arbitrary u_t' can be sent to the system
- We can extract y_t^* from y_t and record it
- We fake the controller by authenticating y_t^* with the acquired watermark

Validating the Attack in Matlab/Simulink

A cyber-physical attack is started at time $t = 700s$







- The attacker disrupt system dynamics (the controller does not perceive it)
- The attack is **not detected**

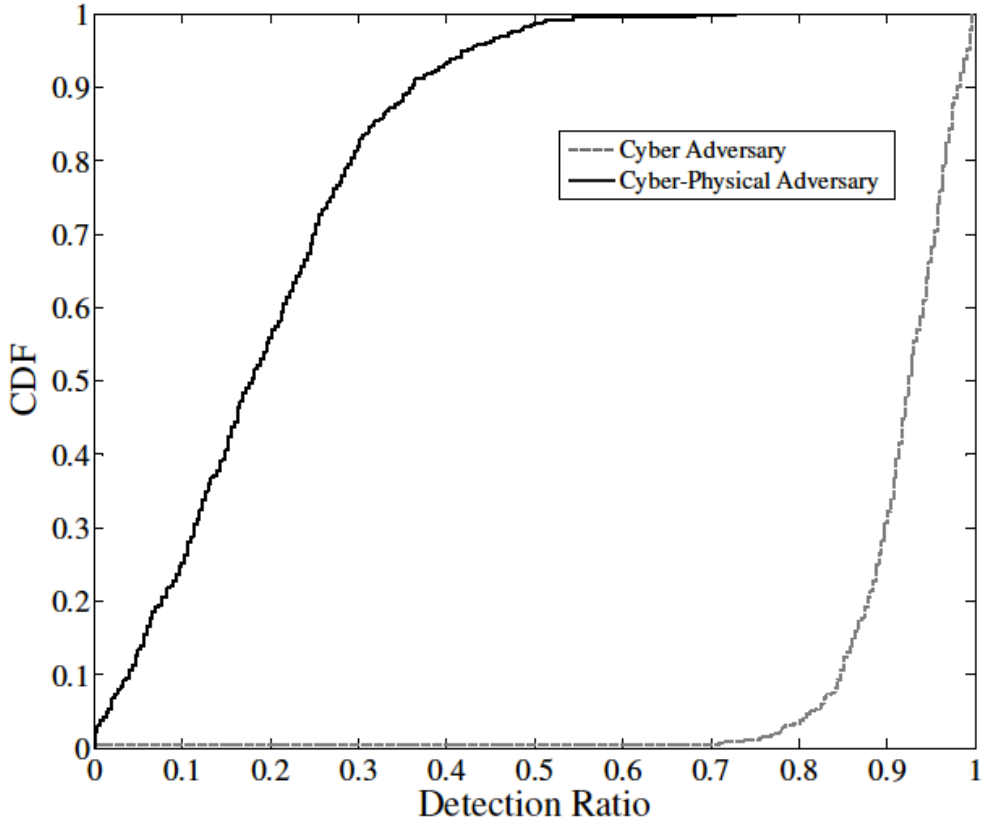
Detection Ratio

To quantify the detector performance we use the **Detection Ratio (DR)** index:

$$DR = \frac{\sum_{t=T_0}^{T_0+T_a} \mathbf{1}_{g_t \geq \gamma}}{T_a} \in [0, 1] \tag{2}$$

- In words: amount of time an attack is detected ($g_t \geq \gamma$) divided by the attack duration T_a
- $DR = 0$ if the attack is never detected, $DR = 1$ if it is always detected

Comparing Cyber and Cyber-Physical Adversary DR



500 simulations

Cyber adversary:
DR average = 0.9

Cyber-physical adversary:
DR average = 0.2

The Watermark-Detector protection scheme by Mo et. al is not sufficiently robust against cyber-physical adversaries

Revisiting the Mo et al. Approach

Towards a multi-watermark cheme

Switch among N watermarks to increase the Detection Ratio

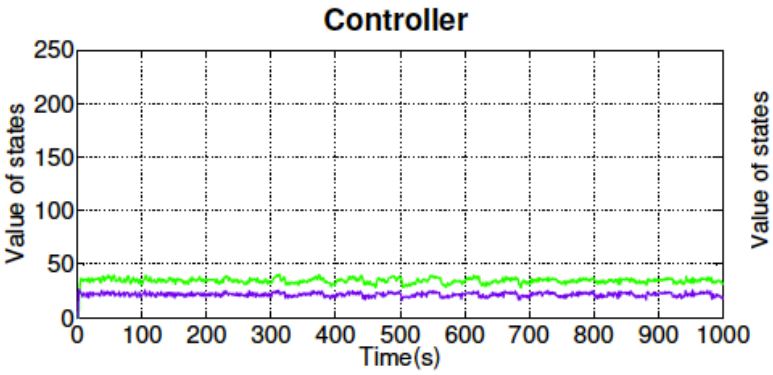
$$\Delta u_t = \Delta u_t^{(s(t,T))} \tag{3}$$

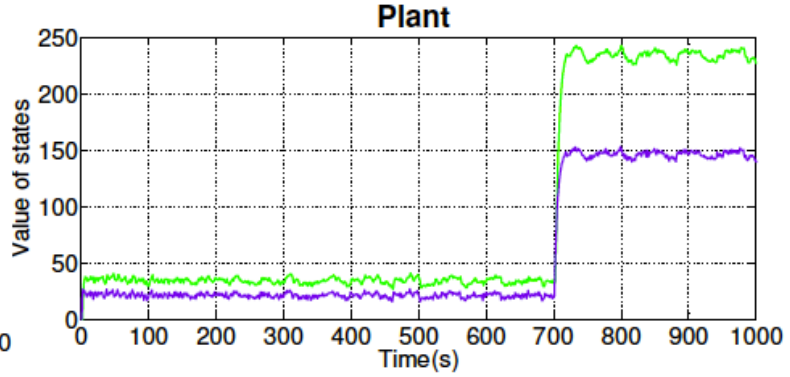
where $\Delta u_t^{(i)}$ with $i \in \mathcal{I} = 0, \ldots, N-1$ is the i-th of the N watermarks, T is the periodicity and T is the switching period

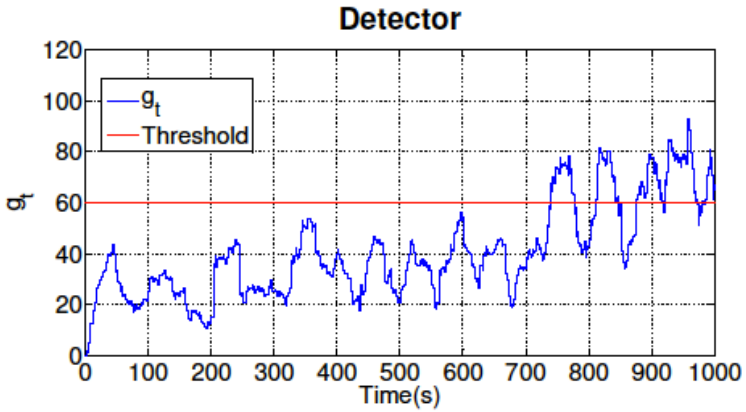
We moved from a static watermark mechanism to an **adaptive** watermark mechanism

Three Watermarks, period T=20s

A Cyber-physical adversary which starts the attack at time $t = 700s$



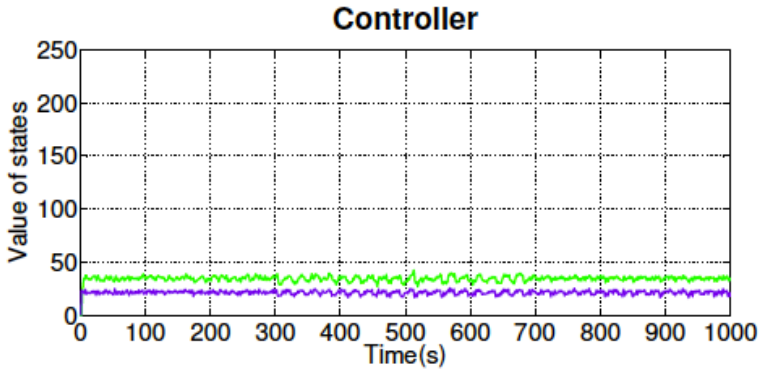


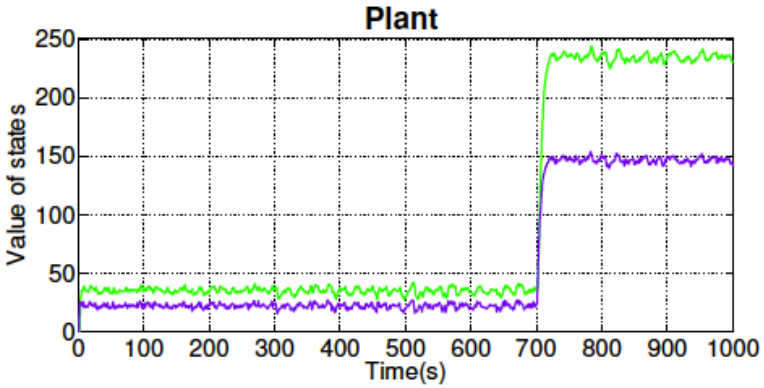


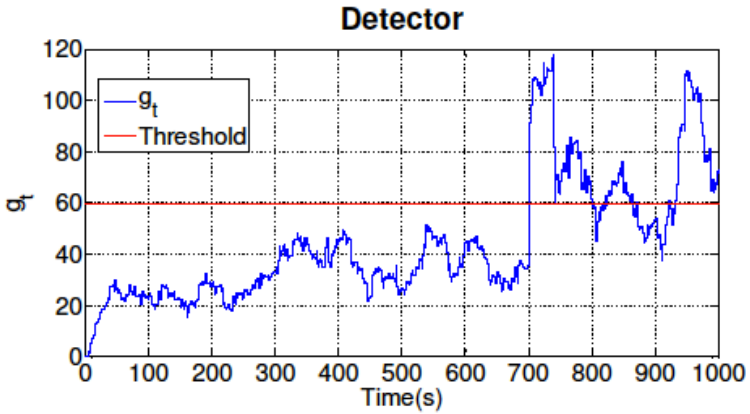
- The attacker disrupt system dynamics (the controller does not perceive it)
- The attack is detected after a while
- Slightly larger oscillations in state dynamics due to the new watermark

Three Watermarks, period T=7s

A Cyber-physical adversary which starts the attack at time $t = 700s$

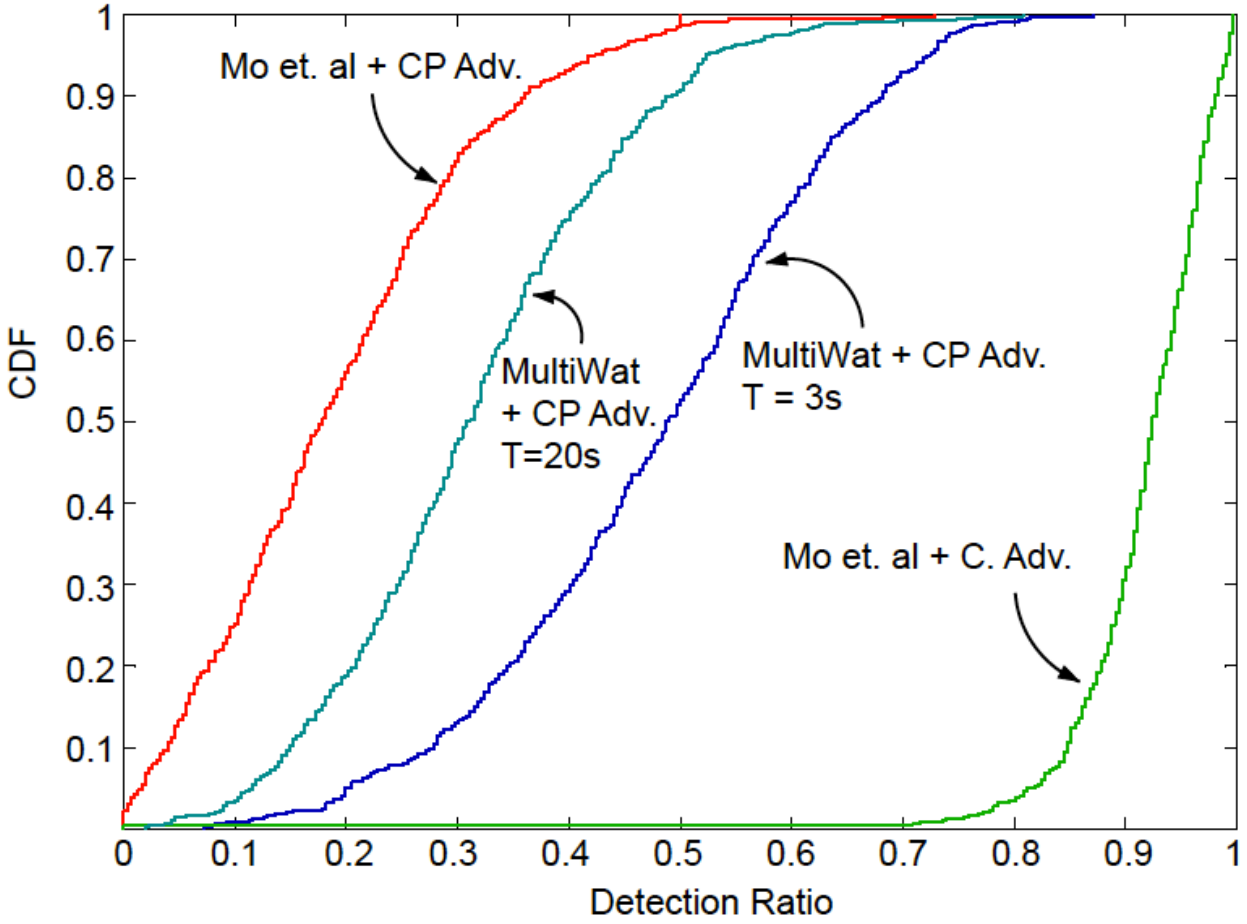






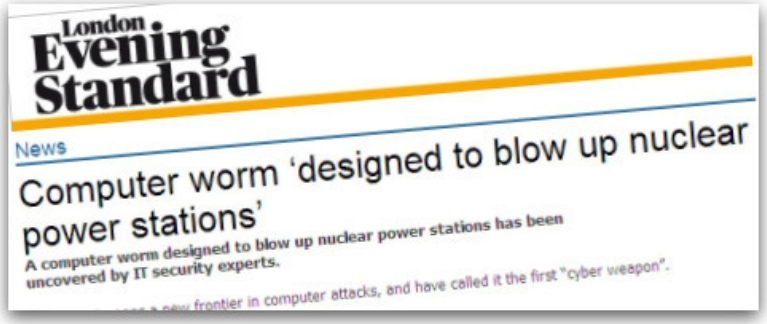
- The attacker disrupt system dynamics (the controller does not perceive it)
- The attack is detected very quickly (DR higher wrt the previous case)
- High frequency oscillations in state dynamics due to the new watermark

Detection Ration vs. Switching Frequency (CDF)



Back to the Suggested Methodology

- Malware moving from IT Systems to Operational Systems
- Wrong configurations, lack of encryption, legacy (vulnerable) systems, third party access, ...


London Evening Standard

News

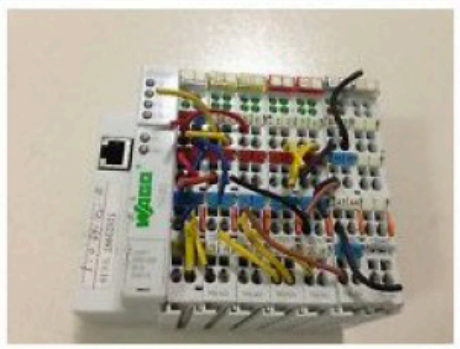
Computer worm 'designed to blow up nuclear power stations'

A computer worm designed to blow up nuclear power stations has been uncovered by IT security experts.

Proposed Methodology

- Foster new theoretical models,
- simulate/emulate case scenarios,
- & **validate results using some testbeds**.

Preparing the Testbeds



WAGO I/O system 750-842 750-402 750-404 750-559 -750-600

From China

Top-rated seller

$400.00

Buy It Now

Free shipping

http://j.mp/1vGPIVp

Siemens S7 300 PLC Trainer, 8 inputs 8 outputs USB/MPI

Learn how to program Siemens PLC's, NO Software

Top-rated seller

$499.99

Buy It Now

http://j.mp/1qViIsG

LEGO Mindstorms EV3 Intelligent Brick # 95646c01

Brand New

Top Rated Plus

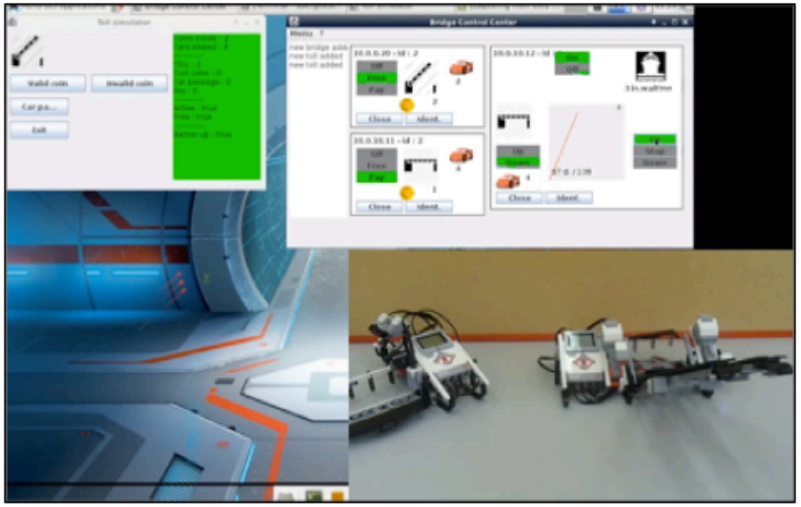
$159.98

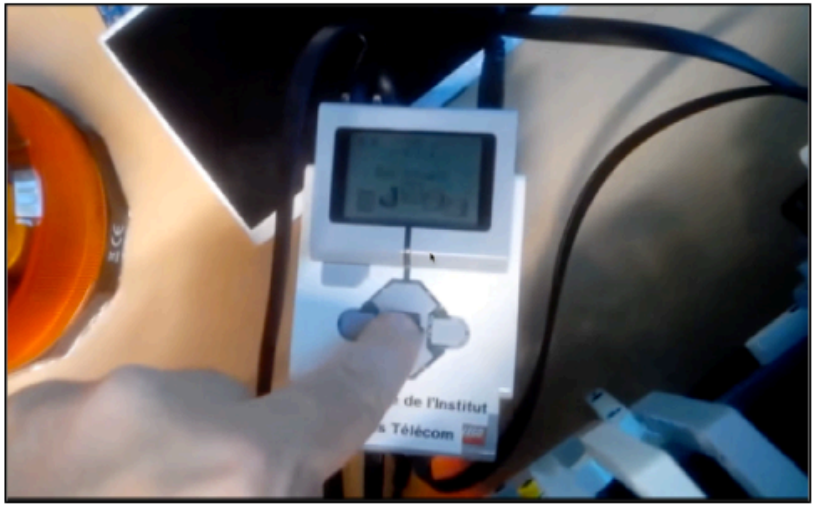
Buy It Now

36 Watchers

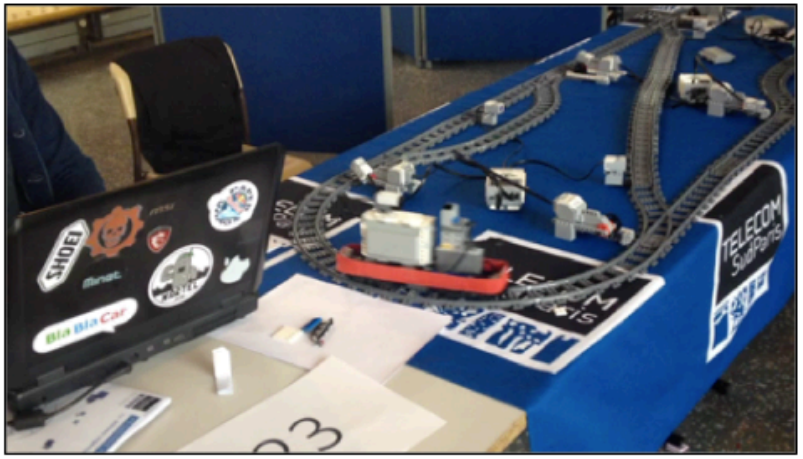
http://j.mp/1lEAxDP

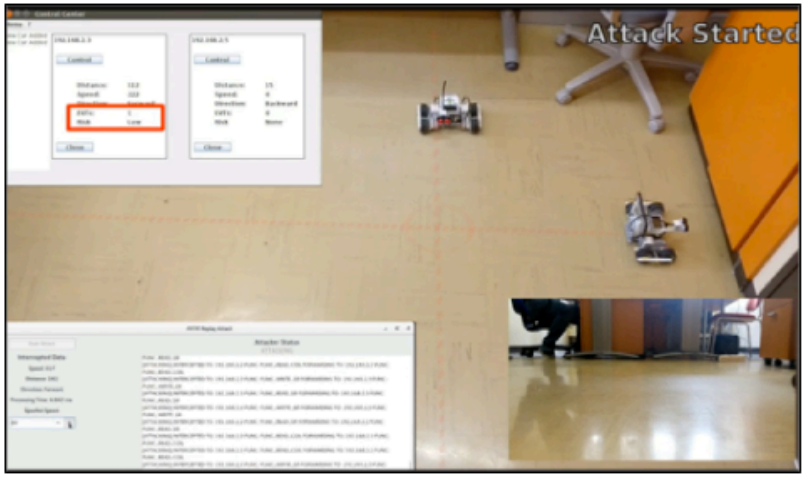
Sample Testbeds



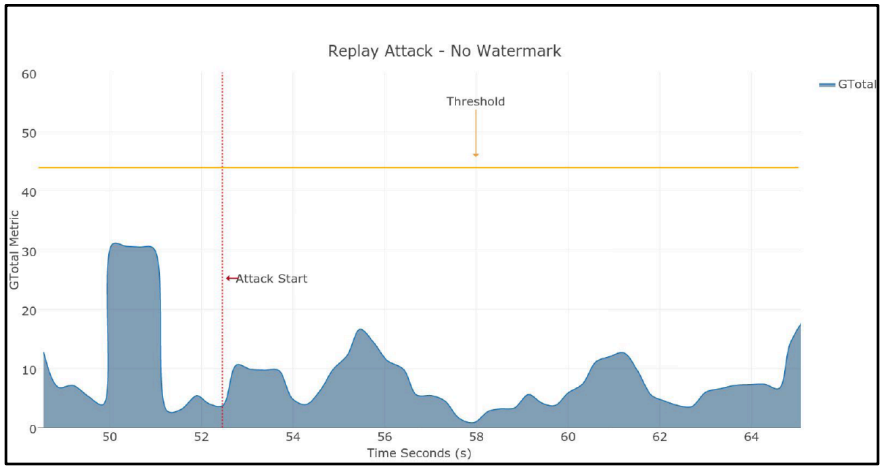


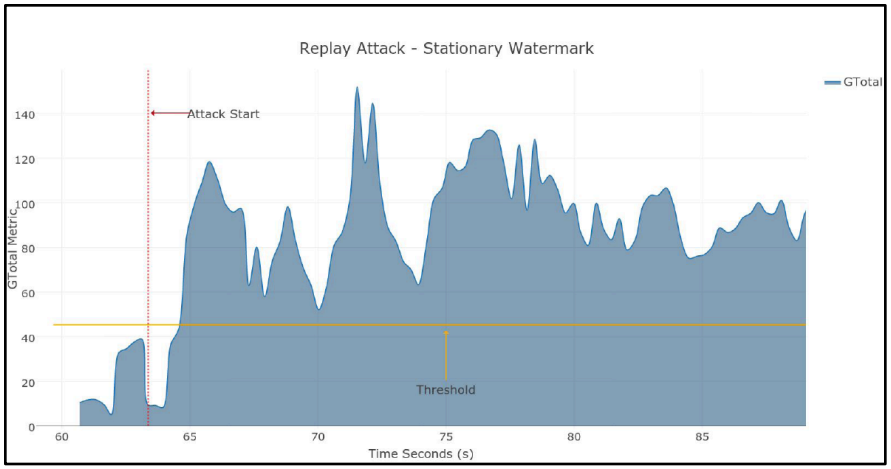
http://j.mp/TSPScada

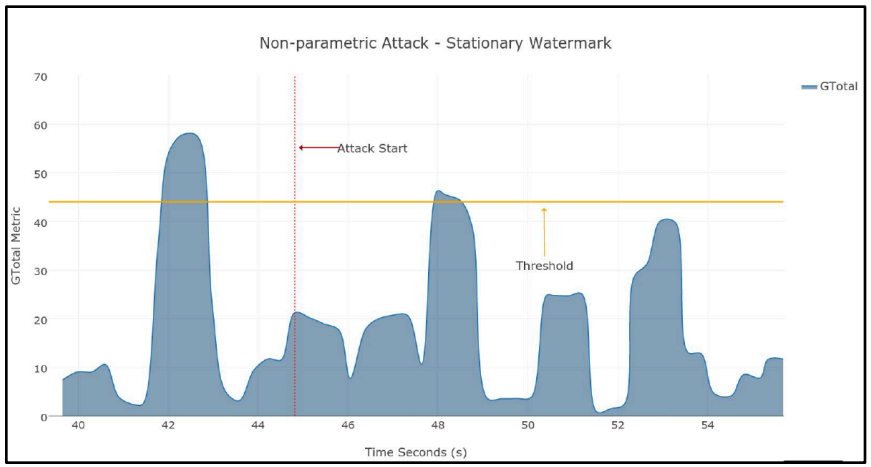


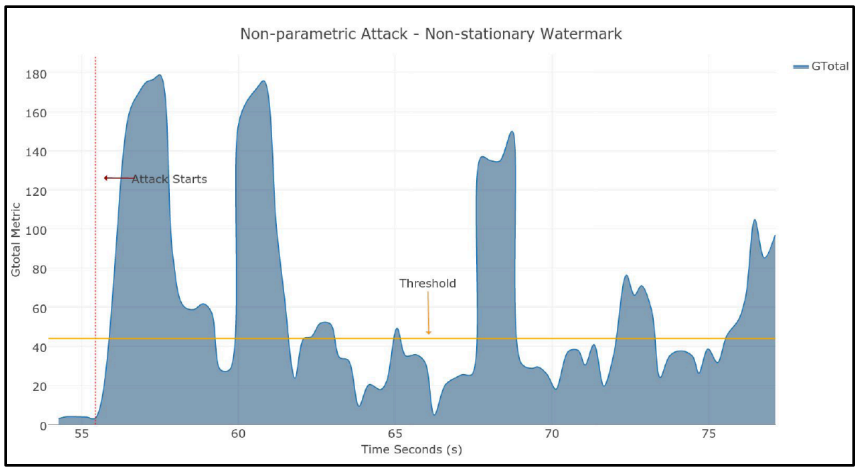


Testbed (*ongoing*) Results

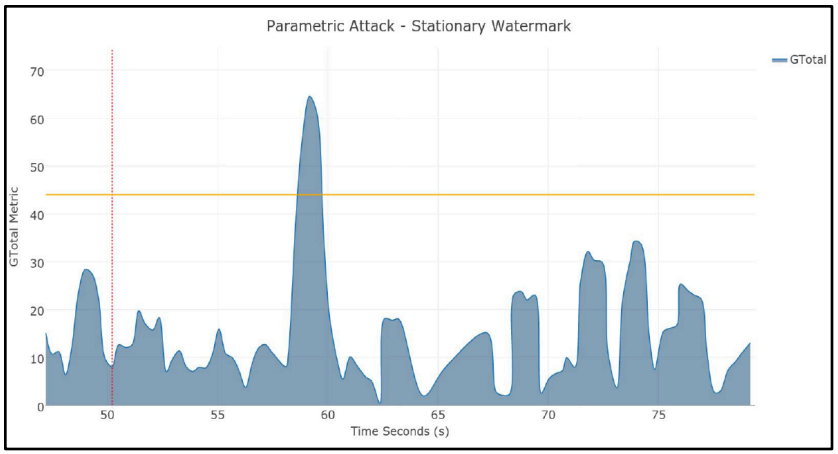


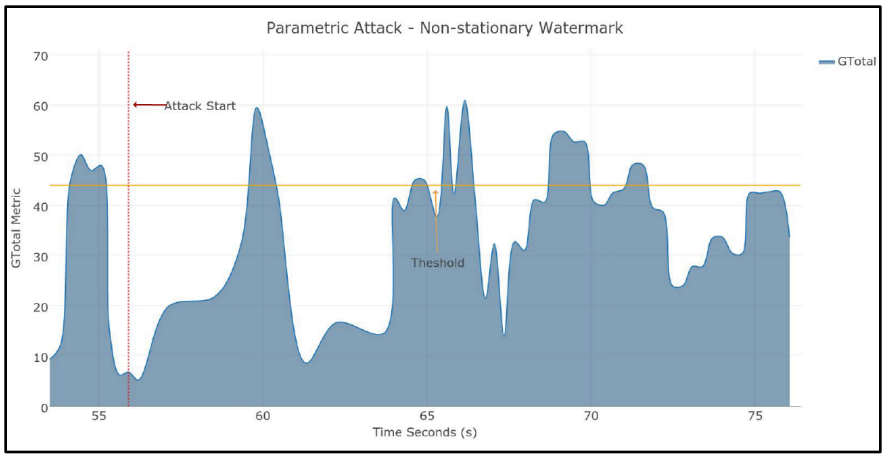






Testbed (*ongoing*) Results





	Replay Attack	FIR Adaptive Attack (Non-parametric Attack)	Parametric Attack
True Positives	35.94%	14.80%	11.37%
False Negatives	64.06%	85.20%	88.63%
False Positives	0.98%	1.66%	1.35%

Outline

- **Brief Introduction**
- **Cyber-Physical Systems**
- **Feedback Control Verification**
- ***Summary & Perspectives***

Summary

- Challenging, multidisciplinary topic
  - Dynamic (networked-control) systems & analysis of large datasets

- Risk Assessment
  - Traditional IT-based methods may still be applicable
  - However, they cannot solve the problem completely
    - Fundamental differences between IT systems & CPSs

- Modeling, from a control-theoretic perspective, is a *must*
  - Pay attention to adversary strategies from the attacker's angle
  - Assume attackers with knowledge about information systems & physical systems at the same time

  - Testbeds for the evaluation of emerging theories, methods & techniques
  - Use practical & real-time environments

Thank You. Questions?

References

- Hirschmann. Why is Cyber Security Still a Problem? *TOFINO Security Series, 2010*.
- Kim & Kumar. Cyber–Physical Systems: A Perspective at the Centennial. *Proceedings of the IEEE*, Vol. 100, pages 1287-1308, May 2012.
- Krotofil & Larsen. Hacking Chemical Plants for Competition and Extortion, *DefCon23, 2015*.
- Texeira et al. A secure control framework for resource-limited adversaries. *Automatica*, 51(1):135-148, 2015.
- Wu, Sun & Chen. A survey on the security of cyber-physical systems. *Control Theory and Technology*, 14(1):2–10, February 2016.
- Rubio, De Cicco, & Garcia-Alfaro. Revisiting a Watermark-based Detection Scheme to Handle Cyber-Physical Attacks. *ARES 2016*, (***Best Paper Runner-Up Award***), August 2016.
- Mo, Weerakkody & Sinopoli. Physical Authentication of Control Systems. *IEEE Control Systems*, Vol. 35, pages 93–109, 2015.